

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2005, 10:44:18 ; Search time 3858 Seconds
(without alignment)
10248.693 Million cell updates/sec

Title: US-10-720-018-1

Sequence: 1 atgactgactcattccac.....atggtggttcactcttcg 816

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sbs:*
12: gb_ay:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	816	100.0	816	AJ583159	AJ583159 Ambrosioz
2	332.8	40.8	1166	CTARDH	U00675 Candida tro
3	305.6	37.5	1614	YSAARDH	L16227 Candida alb
4	275.2	33.7	110000	CR382133_10	Continuation (11 o
5	268.4	32.9	1157	PSARDH	Z46866 P. scriptis
6	238.2	29.2	696	AR345028	AR345028 Sequence
7	228.2	29.2	696	AR345028	AR345028 Sequence
8	218.6	26.8	846	CR382121_09	Continuation (10 o
9	155.8	19.1	110000	CR382133_03	Continuation (4 of
10	130.8	16.0	110000	CR382133_11	Continuation (12 o
11	92.6	11.3	110000	CR382133_06	Continuation (7 of
12	91.4	11.2	110000	AR002134	AF002134 Candida a
13	87.2	10.7	4293	AC105434	AC105434 Magnapor
14	86.8	10.6	126105	AR547359	AR547359 Sequence
15	86.6	10.6	927	AY387591	AY387591 Emergicell
16	81.8	10.0	1485	AY434691	AY434691 Emergicell
17	81.8	10.0	15779	CR382125_21	Continuation (22 o
18	78.6	9.6	110000	CR382131_06	Continuation (7 of
19	78.6	9.6	110000	CR382131_06	Continuation (7 of

C	20	72	8.8	110000	1	AE016827_14	Continuation (15 o
C	21	70.8	8.7	110000	8	CR382133_10	Continuation (11 o
C	22	70	8.6	1063	8	UF081790	U81790 Uromyces fa
C	23	68.6	8.4	409	11	KLAJ9366	AL529366 Kluverom
C	24	68.2	8.4	340900	1	SME591791	AL591791 S. pombe chr
C	25	67.6	8.3	32814	8	SPAC22A12	Z99295 S. pombe chr
C	26	64	7.8	6430	8	ARU320167	AJ320167 Aspergill
C	27	64	7.8	110000	8	CR382138_00	CR382138 Debaryomy
C	28	63.6	7.8	10126	1	AE001711	AE001711 Thermotog
C	29	62.4	7.6	309135	1	AP005946	AP005946 Bradyrhiz
C	30	61.8	7.6	873	6	AR547523	AR547523 Sequence
C	31	61.4	7.5	110000	6	CR382132_02	Continuation (3 of
C	32	61	7.5	786	8	STU245634	AJ245634 Solanum t
C	33	59.8	7.3	822	6	AX772916	AX772916 Sequence
C	34	59.8	7.3	1129	8	DURTROPR1	L20473 Datura stra
C	35	59.4	7.3	15350	8	SPAC8B11	AL021817 S. pombe c
C	36	59	7.2	774	6	E30144	E30144 Arabiditol
C	37	59	7.2	2255	6	E30145	E30145 Arabiditol
C	38	57.8	7.1	306787	1	AE016744	AE016744 Staphyloc
C	39	57.2	7.0	37969	8	CR382137_20	Continuation (21 o
C	40	57.2	7.0	110000	8	CR382137_19	Continuation (20 o
C	41	56.6	6.9	976	8	STU400815	AJ400815 Solanum t
C	42	56.4	6.9	789	8	BT010425	BT010425 Arabidops
C	43	56.4	6.9	825	6	AX772918	AX772918 Sequence
C	44	56.4	6.9	909	8	AK175221	AK175221 Arabidops
C	45	56.4	6.9	1099	8	D88156	D88156 Hyoscyamus

ALIGNMENTS

RESULT 1
LOCUS AJ583159
DEFINITION Ambrosiozyma monospora partial mRNA for NADH L-xylose reductase (ALX1 gene).
ACCESSION AJ583159
VERSION AJ583159.1 GI:42820317
KEYWORDS alx1 gene; NADH L-xylose reductase.
SOURCE Ambrosiozyma monospora
ORGANISM Ambrosiozyma monospora
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycopsidaceae; Ambrosiozyma.
REFERENCE 1 Verho, R., Putkonen, M., Jondesborough, J., Penttila, M. and Richard, P. A Novel NADH-linked L-Xylose Reductase in the L-Arabinose Catabolic Pathway of Yeast J. Biol. Chem. 279 (15), 14746-14751 (2004)
JOURNAL PUBMED 14736891
AUTHORS Richard, P.
REFERENCE 2 (bases 1 to 816)
TITLE Direct Submission
JOURNAL Submitted (18-SEP-2003) Richard P., VTT Biotechnology, VTT, PO box 1500, Tietoele 2, Espoo, 02044 VTT, FINLAND
FEATURES
source
1..816
/organism="Ambrosiozyma monospora"
/mol_type="mRNA"
/strain="NRRL Y-1484"
/db_xref="taxon:43982"
1..816
/gene="alx1"
1..>816
/gene="alx1"
/codon_start=1
/product="NADH L-xylose reductase"
/protein_id="CAE47547.1"
/translation="MTDYIPTPRFDGHLITVGCAGGIALALIKGLAAGSDALADI
DORSTAKOAEYHAKVATDELKKEVPMGSAVDISPDVHVFPQVAAADPKPLPH
LVNTAGCERFPCEPDYPAKAEKMGVYNVLGSLYVSGAPAKPLIKGSIKASAVYLG
MSGAVINDPQVYVNMKAGVTHLAKTLACAEWAKNIRVNSLNPGLTIGPLTKNVIN
GNELYNRWISGIPOQRMSEPKYIGAVLYLSESAASYTTGASLLVDGFTSW"

ORIGIN

Query Match 100.0%; Score 816; DB 8; Length 816;

Best Local Similarity 100.0%; Pred. No. 6,3e-196; Matches 816; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACTGACCTTCACTTCACTTGAATGATGAGCCCTTACCATTTGTCAGAGTGCC 60
 DB 1 ATGACTGACCTTCACTTCACTTGAATGATGAGCCCTTACCATTTGTCAGAGTGCC 60
 QY 61 TGTGGTGGTTAGCTGAGCTTTATCAAGGGTTTGTGGCTACGGTTCTGACATTTGCT 120
 DB 61 TGTGGTGGTTAGCTGAGCTTTATCAAGGGTTTGTGGCTACGGTTCTGACATTTGCT 120
 QY 121 TTGCTGATATGACCAAGAAAGAAGCTGCTCCCAACAGCCGAATATCCAAATATGCGT 180
 DB 121 TTGCTGATATGACCAAGAAAGAAGCTGCTCCCAACAGCCGAATATCCAAATATGCGT 180
 QY 181 ACTGAAAGATGGAAGTTGAAAGAGTTCCAAAGATGGGTTGATATGCTGATATTTCT 240
 DB 181 ACTGAAAGATGGAAGTTGAAAGAGTTCCAAAGATGGGTTGATATGCTGATATTTCT 240
 QY 241 GATTCTGATACCGTTCAACAAGGTGTTGCTCAAGTTGCTAAGATTTTGTAAAGTTGCCA 300
 DB 241 GATTCTGATACCGTTCAACAAGGTGTTGCTCAAGTTGCTAAGATTTTGTAAAGTTGCCA 300
 QY 301 TTGCACTTGGTTAACAAGCTGCTGTTACTGTGAAAACCTTCCATGTGAAGATTTACCCAGCC 360
 DB 301 TTGCACTTGGTTAACAAGCTGCTGTTACTGTGAAAACCTTCCATGTGAAGATTTACCCAGCC 360
 QY 361 AAGAACGCTGGAAGATGGTGAAGGTTAACTGTGGGTTCTTTGTATGTTTCTCAAGCC 420
 DB 361 AAGAACGCTGGAAGATGGTGAAGGTTAACTGTGGGTTCTTTGTATGTTTCTCAAGCC 420
 QY 421 TTGCTAAGCCATTTGATCAAGAGGTATCAAGGGTCTTCTGTGTTTGAATTTGCTTCT 480
 DB 421 TTGCTAAGCCATTTGATCAAGAGGTATCAAGGGTCTTCTGTGTTTGAATTTGCTTCT 480
 QY 481 ATGCTGATGCTGATGTCAGATCTCTCAAAACCAAGTTGTCTAACAATGTCGAAGCT 540
 DB 481 ATGCTGATGCTGATGTCAGATCTCTCAAAACCAAGTTGTCTAACAATGTCGAAGCT 540
 QY 541 GGTGTTATCCATTTGGCTTGAAGCTTTGGCTTGAATGGGCTTAAGTACCAATCAGAGTT 600
 DB 541 GGTGTTATCCATTTGGCTTGAAGCTTTGGCTTGAATGGGCTTAAGTACCAATCAGAGTT 600
 QY 601 AATTCTTTAAACCCAGGTATCACTACGCTCTTGAACCAAGATGTTATCAATGCTAAC 660
 DB 601 AATTCTTTAAACCCAGGTATCACTACGCTCTTGAACCAAGATGTTATCAATGCTAAC 660
 QY 661 GAAGAAATTTGACCAAGATGATCTCTGATATCCCAACAAAGATGTCGAACCAAG 720
 DB 661 GAAGAAATTTGACCAAGATGATCTCTGATATCCCAACAAAGATGTCGAACCAAG 720
 QY 721 GAATACATTTGAGTGGCTTTTGAATGCTGTTGATCTGATCTGATACCTACTGCT 780
 DB 721 GAATACATTTGAGTGGCTTTTGAATGCTGTTGATCTGATCTGATACCTACTGCT 780
 QY 781 GCCAGCTTACTGGTATGATGGTTCACCTTCTTG 816
 DB 781 GCCAGCTTACTGGTATGATGGTTCACCTTCTTG 816

RESULT 2

CTARDH 1166 bp DNA linear PLN 16-MAY-1995
 LOCUS Candida tropicalis D-arabinitol dehydrogenase (ARD) gene, complete
 DEFINITION
 ACCESSION U00675.1 GI:392785
 VERSION U00675.1
 KEYWORDS
 SOURCE Candida tropicalis
 ORGANISM Candida tropicalis

REFERENCE

 JOURNAL MURRAY, J.S., WONG, M.L., MIYADA, C.G., SWICHENKO, A.C., GOODMAN, T.C.
 and WONG, B.
 Isolation, characterization and expression of the gene that encodes
 D-arabinitol dehydrogenase in Candida tropicalis
 Gene 155 (1), 123-128 (1995)
 MEDLINE 95212917
 PUBMED 7598655

TITLE

 JOURNAL MURRAY, J.S., WONG, M.L., MIYADA, C.G., SWICHENKO, A.C., GOODMAN, T.C.
 and WONG, B.
 Isolation, characterization and expression of the gene that encodes
 D-arabinitol dehydrogenase in Candida tropicalis
 Gene 155 (1), 123-128 (1995)
 MEDLINE 95212917
 PUBMED 7598655

AUTHORS

 JOURNAL MURRAY, J.S., WONG, M.L., MIYADA, C.G., SWICHENKO, A.C., GOODMAN, T.C.
 and WONG, B.
 Isolation, characterization and expression of the gene that encodes
 D-arabinitol dehydrogenase in Candida tropicalis
 Gene 155 (1), 123-128 (1995)
 MEDLINE 95212917
 PUBMED 7598655

FEATURES

source

 1. 1166
 /organism="Candida tropicalis"
 /mol_type="genomic DNA"
 /strain="ATCC 750"
 /db_xref="taxon:5482"
 /clone="pCT1A"
 /gene_1db="genomic DNA library in pAB107 from V.F. Kalb
 and J.C. Loper"
 241..1089
 /gene="ARD"
 241..1089
 /gene="ARD"
 /function="catalyzes D-arabinitol + NAD <=> D-ribulose +
 NADH"
 /codon_start=1
 /evidence="experimental"
 /product="D-arabinitol dehydrogenase"
 /protein_id="AA6355.1"
 /db_xref="GI:392785"
 /translation="MDSSTYSYSDNIVPSRLDQVITITGSGSLSAVSRALLAG
 ADIALIDMNLERTQAAADVLFQWGEEDMKHESPIQVSNMNSCIGDAEVEITFLA
 INEHGKAVSLINTAGYAENFPAEYPAKRAENIMKVNIGSGSPYVQAQFAPLQNN
 MTSIIILIGMSGTIVNDPOPCQNNKSGVILHSLACEMAKYIRIVATLSPGYI
 LPTIRIVISHTMKETWESKIPMKMAEKERVGSILVLAISASASYTGHNLVVD
 GYECW"
 ORIGIN

gene

CDS

 Query Match 40.8%; Score 332.8; DB 8; Length 1166;
 Best Local Similarity 65.0%; Pred. No. 1.7e-73;
 Matches 528; Conservative 0; Mismatches 272; Indels 12; Gaps 2;
 QY 14 TTCCAACCTTTAGATTGATGAGCCCACTTAACCATTTGTCAGAGTGCTGTGCTTAG 73
 DB 278 TTCCAAGTTTATGATTGATGAGAAATTAATTAATTAATCACTGCTGTGCTTAG 337
 QY 74 CTGAAGCTTTATCAAGGTTTGTGGCTTACGATGCTTGTGCTGATATCG 133
 DB 338 CCGCTGTGTTTCCAGACCTTTGTAAGAAAGGCTGATATGCTTGTGATATGA 397
 QY 134 ACCAAGAAAGACTGCTCCAAACCAAGCCGATACCAAAATACGCTACGAAGATGA 193
 DB 398 ACTTGAAGAAAGCCCAACAGCCGCAAGAGATGTTTACATGGGCTGAAGAACAAATGA 457
 QY 194 AGTGAAGA-----AGTTCCAAAGATGGTTCATATAGCTGTGATATTTCTGATT 244
 DB 458 AAGAAAGACAGAAATCTCAATTTGTCAGATGATGCTGTGCTGTGATATGATG 517
 QY 245 CTGATACCGTTCAAGGTTTGTGCTCAAGTTGCTAAGATTTTGTGATAGTTGCCATTGC 304
 DB 518 CTGAAGCTGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 577
 QY 305 ACTTGGTTAACAAGCTGTTACTGTGAATACTTCCATGTAAGATTAACCAAGCAAGA 364
 DB 578 TTGTGATTAACAAGCTGTTACTGTGAATACTTCCATGTAAGATTAACCAAGCAAGA 637
 QY 365 ACGCTGAGAAAGATGATGAAGTTAACTGTGGGTTCTTTGATGTTTCTCAAGCCTTGG 424

```

Db      638 ATCCGAAATATTCATGAAGTCATGTTAGATCTTTACGCTTCAACAGCTTTCG 697
Qy      425 CTAAAGCATATGATCAAGAGGATATCAAGGGTCTTCTGTGTTTGAATGTTCTATGT 484
Db      698 CCAAGCCTTGAATCAAAA---TAACTATACCTGGCTATATCTTGAATGTTCTATGT 754
Qy      485 CTGGTCCATTTGTCACATCTCTCAAAACCAAGTTGTCTTACACATGTCACAGGCTGTG 544
Db      755 CTGGTACCATTTGTCAACGATTCACCAACCAATGATATGACCAATGTCCAAGCTGGTG 814
Qy      545 TTATCATTTGGCTAAGCTTGGCTGTGATGGCTTAAATGCAACATCAAGTTAAT 604
Db      815 TCATTCATCTGGCCAGATTTTGAATGAAATGGCCAAATACCAACATCAAGTCAAA 874
Qy      605 CTTTAAACCCAGGTATACATCTACGGTCTCTTGAACCAAGATTTATCAATGTTAAAG 664
Db      875 CTTTGTCCCGAGGTATATTTTGAATCTTCACTTCAACAAAGATTTATGATGTTCACTG 934
Qy      665 AATTGTACACAGATGATCTGTGTATCCCAACAAAGATGTCGAAACCAAGAT 724
Db      935 AATATGAACAGAAATGGAAATCCAGATTCATGAAGAAATGGCCGACCAAAAGAT 994
Qy      725 ACATGGGCTGTTTGAATCTTCTGATCTGCTGCTTCACTACTACTGTTGCA 784
Db      995 TTGCGGTTCTAATTTTGAATCTTGGCTTCTGATTCGCTCTTCACTACTGTTGCA 1054
Qy      785 GCTTACTGTTGATGTTGTTTCACTTCTTGG 816
Db      1055 ATTGCTAGTGAATGTTGTTTCACTTGG 1086

```

RESULT 3
YSAADH 1614 bp DNA linear PLN 21-JUN-1995
LOCUS Candida albicans D-arabinitol dehydrogenase (ArdH) gene, complete cds.

ACCESSION L16227
VERSION L16227.1 GI:295567
KEYWORDS ArdH gene; D-arabinitol dehydrogenase.
SOURCE Candida albicans
ORGANISM Candida albicans

REFERENCE 1
AUTHORS Wong, B., Leeson, S., Grindle, S., Magee, B., Brooks, E. and Magee, P.T.
TITLE Analysis of mutants lacking D-arabitol dehydrogenase and D-arabitol metabolism in Candida albicans: studies of the biosynthetic pathway and the gene that encodes NAD-dependent D-arabitol dehydrogenase
J. Bacteriol. 175 (19), 6314-6320 (1993)

JOURNAL MEDLINE 94012494
PUBMED 8407803
REFERENCE 2 (sites)
AUTHORS Wong, B., Leeson, S., Grindle, S., Magee, B., Brooks, E. and Magee, P.T.
TITLE Analysis of mutants lacking D-arabitol dehydrogenase and D-arabitol metabolism in Candida albicans: construction and analysis of mutants lacking D-arabitol dehydrogenase
J. Bacteriol. 177 (11), 2971-2976 (1995)

JOURNAL MEDLINE 95286472
PUBMED 7768790
REFERENCE 3 (bases 1 to 1614)
AUTHORS Wong, B.
TITLE Direct Submission
SUBMITTED (18-MAY-1993) Brian Wong, Department of Internal Medicine, University of Cincinnati College of Medicine, OH 45267-0560, USA
Original source text: Candida albicans (strain WO1) DNA.
Location/Qualifiers
1..1614

COMMENT
FEATURES
Source
/organism="Candida albicans"
/mol_type="genomic DNA"
/strain="WO1"
/db_xref="taxon:5476"
/map="6; segment O"

```

gene      /cbase_11b="PEMBLYE23 from P.T. Magee"
          /35..1280
          /gene="ArdH"
          /35..1280
          /gene="ArdH"
          /function="catalyzes D-arabinitol + NAD = D-ribulose + NADH"
          /codon_start=1
          /evidence=experimental
          /transl_table=12
          /product="D-arabinitol dehydrogenase"
          /protein_id="AAC37430.1"
          /db_xref="GI:295568"
          /translation="MDSAYMSYDINIVSFRDGLKLVITGSGGILAAVVSRLAKGA
          DVALVDNMLERTQOAAADVLQMGSEOMKXTESPIGVASMSCNIGDAEADLPKAI
          NEHGKSSVAVNTAGYENPFAEYEPKAMLEMKVGLSFPVSGAPLIDNM
          TGSILIGSMGCTIVNDPOCMYTMASGATGILARSLACSMKATNIVATLSPEYIL
          TPLIRNVISGHTMEKTESKIPMKRMAEPFVGSILYLASESASITTGHNLVVDG
          GYECW"

```

ORIGIN

Query Match 37.5%; Score 305.6; DB 8; Length 1614;
Beef Local Similarity 62.9%; Pred. No. 1.3e-66;
Matches 511; Conservative 0; Mismatches 289; Indels 12; Gaps 2;

```

Qy      14 TTCCAACTTTAGATTCGATGCGCCACTTAACATTTGTACAGATGCTGTGGTGTAG 73
Db      469 TCCCAAGTTCGGTTGGATGGAAGAACTAGCATATTAACGGTGGCTCTGGTGGTTGG 528
Qy      74 CTGAAGCTTTAATCAAGGGTTTGTGGCTTACGGTTCTGACATTCCTTCTTATTCG 133
Db      529 CTGCGGTGTATCAAGACTTATATAGCCAAAGGCGCATGTGTCATTATGTCATGA 588
Qy      134 ACCAAGAAAGACTGCTGCCAAACAGCCGAATCCAAATAGCTACTGAAGATGA 193
Db      589 ACTGGAAGAAAGCAACACAGCTGCTAGAGAGCTTACAAATGGGCGCAAGACAAATGA 648
Qy      194 AGTGAAGAAAGTTCCA-----AGATGGGTTCATATGCTGTGATATTTTGTGAT 244
Db      649 AAGGTAATAGATTCACCAATCGGTGAGAGTGTCTGTGATGTAATTTGGCCATG 708
Qy      245 CTGAATCCGTTCAAGAGTGTTCCTCAAGTTGCTAAGATTTTGTGAATGTTGCCATTGC 304
Db      709 CTGAAGCTGTGACTTTCATTTCAAGGCATCAACGAAACACACGGCAAAATCTCAAGTG 768
Qy      305 ACTGTGTTAACAAGCTGTGTTACTGTGAATCTTCCATGTGAAGATTAACCAAGCAAGA 364
Db      769 TCTTGTCAACACTGCCGTTTACGCTGAACCTTCCAGCTGAAGAGTACCACAGCAAGA 828
Qy      365 ACGCTGAAGAGATGATGAGGTTAATCTTGTGGGTTCTTGTATGTTTCTCAAGCTTTCG 424
Db      829 ACGCTGAAGAACTTATGAAATGTTAACGGGTTGGGATCATTTCACTTCCCAAGCTTTCG 888
Qy      425 CTAAAGCATATGATCAAGAGGATATCAAGGGTCTTCTGTGTTTGAATGTTCTATGT 484
Db      889 CTAAAGCATATGATCAAAA---CAACATGACCGGATGATCATTTGATCGGTCATATGT 945
Qy      485 CTGGTCCATTTGTCACATCTCTCAAAACCAAGTTGTCTTACAAATGTCACAGGCTGTG 544
Db      946 CCGGTACATCTGTAAAGACCAACACCAATCATGTACATGATGTCACCAACCCGGTG 1005
Qy      545 TTATCATTTGGCTAAGCTTGGCTGTGATGGGCTTAAATGTAACAATCAAGATTAAT 604
Db      1006 TCATTCATTTAGCCAGATCATGCTGTGATATGAGGCTTAATCAATATCAATGATCAACA 1065
Qy      605 CTTTAAACCCAGGTATACATCTACGGTCTTGAACCAAGATTTATCAATGTTAAAG 664
Db      1066 CATGTGCGCAGGGTATCATTTAACCCTATGACAAAGAGTTATATGAGACACACCG 1125
Qy      665 AATTGTACACAGATGATCTGTGATCTTCCACAAAGAAATGTCGAACCAAGAT 724
Db      1126 AATGAAAGACAGATGGAAATCAAAAGATCCCAATGAAGAAATGGACAAAGAGT 1185

```

QY 725 ACATTTGGTCTGTTTGTACTTCTGTAATCTGCTGCTTCAACACTACTGCTGCA 784
DB 1186 TTGTTGGTTCTTACTTACTTGGCTTCAAGATCTGCTTACTTACTACTGCTGCA 1245
QY 785 GCTTACTGTTGATGTTGCTTCACTTCTTGG 816
DB 1246 ACTTGTGCTTGAAGGGGTTACAGATCTGG 1277

RESULT 4
CR382133_10/c
WPCOMMENT
Sequence split into 13 fragments LOCUS CR382133 Accession CR382133

Fragment Name Begin End
CR382133_00 1 110000
CR382133_01 100001 210000
CR382133_02 200001 310000
CR382133_03 300001 410000
CR382133_04 400001 510000
CR382133_05 500001 610000
CR382133_06 600001 710000
CR382133_07 700001 810000
CR382133_08 800001 910000
CR382133_09 900001 1010000
CR382133_10 1000001 1110000
CR382133_11 1100001 1210000
CR382133_12 1200001 1249565
Continuation (11 of 13) of CR382133 from base 1000001 (CR382133 Debaryomyces hansenii ch

Query Match 33.7%; Score 275.2; DB 8; Length 110000;
Best Local Similarity 60.1%; Pred. No. 6e-59;
Matches 492; Conservative 0; Mismatches 308; Indels 12; Gaps 2;

QY 14 TTCAACTTTTATGATTCAGTCCCTTAACCATTTGTACAGTGTCTGTGTTTAC 73
DB 28160 TTCCAAAGCTTAAAGATTAGTGTAGTATGCTATCTTGACCGTGAATCAGAGGATTGG 28101
QY 74 CTGAAGCTTAAATCAAGGGTTTGGCTCAAGTTTGAACATTTGCTTGTGTTATTCG 133
DB 28100 CTGCTGTGTCTTGAAGCATTTATGCTCAAGGACCAAAATTTGCTTGTATGTA 28041
QY 134 ACCAAGAAAGACTGCTGCAACCAAGCCGAATACCAAAATGCTACTGAAAGTTGA 193
DB 28040 ATTTGAAAGAAACCAAGGCTGACGAAGAAATGGAACATGAGGAGCAAGAAATCATTTGA 27981
QY 194 AGTTGAAAGA-----AGTTCCAAAGATGGTTTCATATGCTGTGATTTTCTGAT 244
DB 27980 AAGGTGCAATGAAAGCCCAAGTGGAGCTGTTTCCGCAATGCTTGTATATATTTGGGANT 27921
QY 245 CTGATACCGTTCAAGAGTGTTCGTCAGATTGTAAGATTTGTGAAGTTGCCATTGC 304
DB 27920 TTGAACAGGTGGAAGATGCTTATGAACATCAATGAACCAACATATGTTGGCTGANT 27861
QY 305 ACTTGGTTTAAACAGCTGTTTACTGTGAACATTTCCATGTGAAGTTACCCAGCCAGA 364
DB 27860 TGTATTAATCAATCTCAGAGTTACTGTGAATTTCCCGGCCGAAGAAATCCCATCTGCCA 27801
QY 365 ACCGTGAGAAAGTGTGAAGTTAACTTGTGGTCTTGTATGTTTCTCAAGCTTTG 424
DB 27800 ATGCTGAAGATTTTGAAGGTCATGATTAAGTGTGATTTCAAGTATCAAGCAATTTG 27741
QY 425 CTAAAGCATTTGATCAAGAAAGGATCAAGGTTCTGCTGTTTGTGATTTGTTCTATGT 484
DB 27740 CCCGTCATTGATCA---GCAGCAATTAAGAAAGATTCATCTTTTATGTTGGTGCATGT 27684
QY 485 CTGTGTCCATTGTCAACGATCTCTCAAAACCAAGTTGTCTTCAACATGTCCAGGCTGTG 544
DB 27683 CTGGAACATTTGTGAACGACCAACCAACATGATGTATGAACATGTCCAGGAGGAG 27624
QY 545 TTATCCATTTGCTAAAGCTTTGGCTTGTGAATGGCTGAAGTCAACATCAAGTTAAT 604
DB 27623 TTATTCATCTTGAACAGATCATTTGGTGTGAATGGGCAAAATTTCAACATCAAGTTAACA 27564

QY 605 CTTTAAACCAAGTTATCATCTACGCTCTTTGACCAAGAAATGTTATCATGTACAGAG 664
DB 27563 CTTTGAATCCGATATATATCTTAACCCCTTTGACCAAGAAATGTTATCTTCTGATTCG 27504
QY 665 AATTGTACACAGATGATCTCTGTATGCCACAAACAAAGTCCGACCAAGGAAT 724
DB 27503 ATATGAAGGAAGCATGGAAATCAAGGTCCTCAATGAAGAAATGGCCAAACAAAGGAAT 27444
QY 725 ACATTTGGTCTGTTTGTACTTCTTCTGAAATCTGCTGCTTCAATCACTACTGTTGCCA 784
DB 27443 TTGTTGGCTGATCTTCTTACTTGTGCGCAGAGCTGCTTCTAGTTTACTACAGTGCACA 27384
QY 785 GCTTACTGTTGATGTTGCTTCACTTCTTGG 816
DB 27383 ACTTAGTTGTGATGTGGCTACAGATGCTGG 27352

RESULT 5
PSARDH
LOCUS
DEFINITION
P.stipitis ARDH gene encoding D-arabitol dehydrogenase.
VERSION
Z46866.1 GI:758241
KEYWORDS
D-arabitol dehydrogenase.
SOURCE
Pichia stipitis
ORGANISM
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetaceae; Pichia.

REFERENCE
AUTHORS
TITLE
Hallenborn, J., Walfridsson, M., Penttila, M., Keranen, S. and
Hahn-Hagerdal, B.
A short-chain dehydrogenase gene from Pichia stipitis having
D-arabinitol dehydrogenase activity
Yeast 11 (9), 839-847 (1995)
MEDLINE
JOURNALS
PUBMED
7483848
2 (bases 1 to 1157)
REFERENCE
AUTHORS
TITLE
JOURNALS
Direct Submission
Submitted (30-NOV-1994) Johan Hallenborn, Applied Microbiology, Lund
Institute of Technology, Chemical center, Getingeaven 60, Lund,
S-221 00, Sweden

FEATURES
source
location/Qualifiers
1..1157
/organism="Pichia stipitis"
/mol_type="mRNA"
/strain="CBS 6054"
/db_xref="taxon:4924"
1..1157
/gene="ARDH"
/feature="ARDH"
191..1027
/function="NAD dependent, oxidises D-arabitol to
D-ribulose"
/codon_start=1
/evidence=experimental
/product="D-arabitol dehydrogenase"
/protein_id="CA86939.1"
/db_xref="GI:763164"
/db_xref="GDB:P50167"
/db_xref="UniProt/Swiss-Prot:P50167"
/translation="MDSYANVVFNFRLDGLAIITGSGGLAAVISRALIAGADVA
LIDNMLERTKAKEVIGMBETLKGEHSAIQGVASWSNIGDAEVDATFSSINER
HGKIALDLINTRAGCENPAPATPATNAESIMKNGSGSFYSGSFAPLIONLRGS
IILGNSGTTVDNPQPCQWYNNMKAGVYIIVRSIACEMAKYNI RNVNLTSPGYLTP
ITRNVISGHTTEKMESEKTPMKRAEPPEFVGSILYLAETASSTTGHNLVVDGVE
CW"

gene
CDS

ORIGIN

Query Match 32.9%; Score 268.4; DB 8; Length 1157;
Best Local Similarity 60.1%; Pred. No. 3.6e-57;
Matches 487; Conservative 0; Mismatches 311; Indels 12; Gaps 2;

QY 5 CTGACTACATCCAACTTTAGATTGATGCGCCTTAACATGTGACAGGTGCTGTG 64
 DB 207 CTAACGGTGTTCCTCACTTCAGATTGAGCGAAGATTGGCTATTATTACCGGAGGTTCTG 266
 QY 65 GTGGTTAGCTGAAGCTTTATCAAGGGTTTGTGGCTTCAAGGTTCTGACATTTGCTTTC 124
 DB 267 GTGGTTAGCTGAAGCTTTATCAAGGGTTTGTGGCTTTCAGGCTTATGTTGCTTCA 326
 QY 125 TTGATATGACCAAGAAAAGACTGTGCTCCAAACAGCCGATATCAACAATACGCTACTG 184
 DB 327 TTGATATGACCAAGAAAAGACTGTGCTCCAAACAGGATTTTGGCTGGGGTGAAG 386
 QY 185 AAGATTGAAGTTGAAAAGAGTTCCAAAGATGGGTTCAATG-----CTGTGATA 235
 DB 387 AGACGTTGAAGGGTGAACACGCTTCAGCCATCGGCCCAAGTTCCGCTGCTGTGACACA 446
 QY 236 TTTGATTTCTGATACCGTTCAAGAGTGTGCTCAAGTTGCTAAGATTTTGGTAAGT 295
 DB 447 TTGGGAGTGTGAGGCAATGACGCTACTTTCAGCTCCATCAACGACACCGCGACAGA 506
 QY 296 TGCATTGCACTTGGTTAACAGAGCTGTGTTACTGTGAAAACCTCCATGTGAAGATTACC 355
 DB 507 TCGCTGACTGTGTGATTAAACACGCTGATATCTGTGAAAACCTCCCTGCGGAAAGTACC 566
 QY 356 CAGCCAGAAGCTGTGAAGAGTGTGAAGTTAACTTGTGGCTTCTGTATGTTTCTC 415
 DB 567 CGGCTACTAGCGCTAAAGCATGAAAGTGAACGGTTTGGCTCATTTCTACGTTTGGC 626
 QY 416 AAGCTTTGCTAAGCATTTGATCAAGAAGTATCAAGGGTGTCTGTGTTTGTGATTG 475
 DB 627 AATCGTTGCTAGACATTTGATGACGAA--CACTTGAGAGGCTCTATCATCTTGATTG 683
 QY 476 GTTCTATGTCTGTGCTCAATGTCAACGATCTCAAAACCAAGTTGCTCAACATGTCCA 535
 DB 684 GCTCATGTCTGTGAACATTTGTCAACGACCCCAACCCCAAGTATGTACACATGTCCA 743
 QY 536 AGGCTGTGTATTCATTTGGCTAAGACTTTGGCTTGTGAATGGCTTAAGTCAACATCA 595
 DB 744 AGGCTGTGTATTCATTTGGCTAAGACTTTGGCTTGTGAATGGCTTAAGTCAACATCA 803
 QY 596 GAGTTAATCTTTAAACCAAGTTATCATCTACGCTCTTGAACCAAGATGTATCAATG 655
 DB 804 GAGTTAATCTTTAAACCAAGTTATCATCTACGCTCTTGAACCAAGATGTATCAATG 863
 QY 656 GTAAAGAAATTTGTACAAAGATGATCTGTGATCTCCCAACAAAGAAATGTCCGAAAC 715
 DB 864 GCCACACAGAGATGAAAGAGCTGTGGAATCCCAAGATCCCAATGAAGATGTCCGAAAC 923
 QY 716 CAAGGAATTAATGTGTCTGTTTGTACTTGTGCTTCTGAATCTGTCTCTTCACTCA 775
 DB 924 CAAGGAATTAATGTGTCTGTTTGTACTTGTGCTTCTGAATCTGTCTCTTCACTCA 983
 QY 776 CTGGTCCAGCTTACGTGTGATGTGTGT 805
 DB 984 CGGCGCACAAATTTGTTGTGACGGAAGAT 1013
 RESULT 6
 AR031556 696 bp DNA linear PAT 29-SEP-1999
 LOCUS AR031556 Sequence 6 from patent US 5866382.
 DEFINITION AR031556
 ACCESSION AR031556
 VERSION AR031556.1 GI:5945845
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 696)
 AUTHORS Hallborn,J., Penttila,M., Ojamo,H., Walfridsson,M., Airaksinen,U.,
 Keranen,S. and Hahn-Hagerdal,B.
 TITLE Xylose utilization by recombinant yeasts
 JOURNAL Patent: US 5866382-A 6 02-FEB-1999;
 FEATURES Location/Qualifiers

source 1. 696
 /organism="unknown"
 /mol_type="unassigned DNA"
 ORIGIN
 Query Match 29.2%; Score 238.2; DB 6; Length 696;
 Best Local Similarity 64.4%; Pred. No. 1.7e-49;
 Matches 373; Conservative 0; Mismatches 203; Indels 3; Gaps 1;
 QY 227 CCTGTGATATTTCTGATTTCTGATTAACGTTCAACAGGTGTTGCTCAAGTGTGAAGATT 286
 DB 107 CTTGCAACATTGGGATGTCTGAGCGATGACGCTACTTTCAGCTCCATCAACGAACACC 166
 QY 287 TTGTGAATGTGCATTGACATTGGTTAAACAGCTGTGTTACTGTGAAAACCTTCCATGTG 346
 DB 167 ACGGCAAGATGCTGACTGTTGATTAAACCGCTGATATCTGTGAAAACCTTCCGCGC 226
 QY 347 AAGATTACCCAGCCCAAGACGCTGAAGATGTGAAGATTAACTTGTGGTCTTGT 406
 DB 227 AAACGTTACCGGCTACTAACGCTGAACCATCATGMAAGTGAACGTTTGGCTCATTTCT 286
 QY 407 ATGTTCTCAAGCCTTGTCTAAGCATGATCAAGAAGTATCAAGGGTGTCTGTG 466
 DB 287 ACGTTCCGAATTCGTTGCTGAGCCATGATGACGAA--CACTTGAGAGGCTCTATCA 343
 QY 467 TTTGATGTGTTCTATGTCTGTGCTCATGTCAACGATCTCAAAACCAAGTTGTCTACA 526
 DB 344 TCTTGAATGGCTCAATGCTGTGAACAATTGTCAACGACCCCAACCCCAATGATGATACA 403
 QY 527 ACATGTCCAAGGCTGTGTATTCATTTGGCTAAGACTTTGGCTTGTGAATGGCTAAGT 586
 DB 404 ACATGTCCAAGGCTGTGTATTCATTTGGCTAAGACTTTGGCTTGTGAATGGCTAAGT 463
 QY 587 ACAACATGAGATTAATCTTTAAACCAAGTTATCATCTACGCTCTTGAACCAAGATG 646
 DB 464 ACAACATGAGATTAATCTTTAAACCAAGTTATCATCTACGCTCTTGAACCAAGATG 523
 QY 647 TTATCAATGTGTAAAGAAATTTGTACACAGATGATCTGTGATCTCCCAACAAAGAA 706
 DB 524 TGATTTCTGGCCACACAGAGATGAAAGAGCTGGGATCCAGATCCCAATGAAGATA 583
 QY 707 TGTCCGAACCAAGAAATTAATGTGTGCTGTTTGTACTTGTCTTCTGAATGTGTGCTT 766
 DB 584 TGGCCGAACCAAGAAATTAATGTGTGCTGTTTGTACTTGTCTTCTGAATGTGTGCTT 643
 QY 767 CATACACTGTGTGCCACTTACTGTGTGATGTGTGT 805
 DB 644 CTAACACTAGCGGCGCACAAATTTGTTGTGACGGAAGAT 682
 RESULT 7
 AR345028 696 bp mRNA linear PAT 17-AUG-2003
 LOCUS AR345028 Sequence 6 from patent US 6582944.
 DEFINITION AR345028
 ACCESSION AR345028
 VERSION AR345028.1 GI:33741148
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 696)
 AUTHORS Hallborn,J., Penttila,M., Ojamo,H., Walfridsson,M., Airaksinen,U.,
 Keranen,S. and Hahn-Hagerdal,B.
 TITLE Production of ethanol from xylose
 JOURNAL Patent: US 6582944-A 6 24-JUN-2003;
 FEATURES Location/Qualifiers
 source 1. 696
 /organism="unknown"
 /mol_type="mRNA"
 ORIGIN
 Query Match 29.2%; Score 238.2; DB 6; Length 696;
 Best Local Similarity 64.4%; Pred. No. 1.7e-49;
 Matches 373; Conservative 0; Mismatches 203; Indels 3; Gaps 1;

Matches 373; Conservative 0; Mismatches 203; Indels 3; Gaps 1;

QY 227 CCTGTGATTTCTGATTCGATACCGTTCAAGAGTTTGTCTCAAGTTCTAAGATT 286
DB 107 CCTCAACATTTGGGAGTACGAGTAGACGCTACTTTCAGCTCCATCAACGACACC 166
QY 287 TTGGTAAGTTGCCATTCGATCTTGGTTAAACAGAGCTGTCTTCTGAAAACTTCCATG 346
DB 167 ACGGCAAGATGGCTGACCTGTTGATTAAACCGCTGAGTACGTGAAAACTTCCGCG 226
QY 347 AAGATTACCCAGCCAGAACGCTGAGAAAGATGTAAGTTAACTTGTGGCTTTTGT 406
DB 227 AAGAGTACCCGCTTACTAACGCTGAAGCATGAAAGTGAACGGTTTGGCTCATCT 286
QY 407 ATGTTCTCAAGCCCTTGTCTAAGCATGATCAAGAGGTATCAAGGGTCTTCTGTG 466
DB 287 ACGTTGCGAATCGTTCCGTAGACCATTTGATCAGAA---CACTTGAGAGGCTTATCA 343
QY 467 TTTTGATTTGTTCTAATGTTCTGTCATTCATTCAGATCTCAAAACCAAGTTGCTACA 526
DB 344 TCTTGATTTGGCTCAATGCTGGAACAAATTTGCAACGACCAACCCCAATGTAATGACA 403
QY 527 AATGTCCAGGCTGTGTATTCATTTGGCTAAGACTTTGGCTTGTGAATGGGCTAAGT 586
DB 404 AATGTCCAGGCTGTGTATTCATTTGGCTAAGACTTTGGCTTGTGAATGGGCTAAGT 463
QY 587 ACAATCATAGATTAACTTTTAAACCAAGTTTACATCTGAGTCTTGTGACCAAGATG 646
DB 464 ACAATCATAGATTAACTTTTAAACCAAGTTTACATCTGAGTCTTGTGACCAAGATG 523
QY 647 TTATCAATGTAAACGAAATTTGTAACAAGATGATCTGTGTAATCCCAACAAAGAA 706
DB 524 TGATTTTGGGCAACAGATGAGAAAGAGCTGGGAATCCAAAGATCCCAATGAAGAA 583
QY 707 TGTCCGAACCAAGAAATACATTTGCTGCTTTTGTACTTGTCTTCTGAATTTGTGCTT 766
DB 584 TGGCCGAACCAAGAAATTTGCTGCTTCTTATTTACTTGTGCAAGCAAGACTGCTTCT 643
QY 767 CATACACTAGTGTGCGAGCTTACTGTTGATGATGATT 805
DB 644 CTTACACTAGTGTGCGAGCTTACTGTTGATGATGATT 682

RESULT 8
AR544898 846 bp DNA linear PAT 08-OCT-2004
DEFINITION AR544898
ACCESSION AR544898
VERSION AR544898.1 GI:53938073
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 846)
Weinstein, K.G. and Bush, D.
Nucleic acid sequences relating to *Candida albicans* for diagnostics
and therapeutics
JOURNAL Patent: US 6747137-A 29 08-JUN-2004;
FEATURES
location/Qualifiers
source
1..846
/organism="Unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 26.8%; Score 218.6; DB 6; Length 846;
Best Local Similarity 62.0%; Pred. No. 1.5e-44;
Matches 385; Conservative 0; Mismatches 224; Indels 12; Gaps 2;

QY 14 TTCGAATTTTATGATTCGATGCGCACTTAACATTTGTCAAGATGCTGTGGGTTTATG 73
DB 218 TTCGAAGCTTCCGTTTGGATGAAACTATGATTAACCGGTGCTCTGGGTTTGG 277
QY 74 CTGAAGCTTTAATCAAGGGTTGTTGGCTTACGCTTTCAGATTCCTTGTGATATG 133

DB 278 CTGCGGTGTATCAAGACTTTTATTCACCAAGGTGCGATGTTGATTTGATGATATGA 337
QY 134 ACCAAGAAAAGATGCTGTGCCAACAAGCCGAATACCAATATAGCTACTGAAAGATTGA 193
DB 338 ACTTGGAAAAGAACAAACAAAGCTGTAGAACGCTTTACATATGGGGGAAGACAAATGA 397
QY 194 AGTTGAAGAAGTTCCAA-----AGATGGTTATATGCTGTGATTTTCTGATT 244
DB 398 AAGTAATATGAAATCACCAATCGTCAAGTGAAGTCTTGTGATTAATATGGCGATG 457
QY 245 CTGATACCTTTCACAAAGTGTGCTCAAGTTGCTAAGTTGCTAAGATTTTGTGAATGGCATTC 304
DB 458 CTGAAGCTGTGACTTTGACATTCAAAGCCATCAACGAAACACAGGCAAAATCTCAAGTG 517
QY 305 ACTTGTTAAACACAGCTGTGTTACTGTGAATACTTCCATGTAAGATTAACCAAGCAGA 364
DB 518 TCTTGTTCACACCTGCGGCTTAACCTGTAATACTTCCAGCTGAAGATTAACCAAGCAGA 577
QY 365 ACGCTGAAGAAGATGTAAGGTTAACTTGTGGTCTTTGTATGTTTCTCAAGCTTTG 424
DB 578 ACGCTGAAGAAGATGTAAGGTTAACTTGTGGTCTTTGTATGTTTCTCAAGCTTTG 637
QY 425 CTGAAGCTATCAAGAAAGATATCAAGGTCCTTCTGTTGTTGATTTGATTTGATTTGAT 484
DB 638 CTGAAGCTATCAAGAAAGATATCAAGGTCCTTCTGTTGTTGATTTGATTTGATTTGAT 694
QY 485 CTGTCGCAATTTGCAAGATCTTCAAAACCAAGTTGTCTACAAATGTCGAAGCTGTG 544
DB 695 CCGTACCATTTGTAAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAG 754
QY 545 TTATCAATTTGGCTAAGACTTTGCTGTGTAATGGGCTAAGATCAATCAAGTATTT 604
DB 755 TCATTCATTTTGGCAAGATCAATTTGCTGTGATGAGGCTTAATATCAATATCAGATCAACA 814
QY 605 CTTTAAACCAAGTTTACATCT 625
DB 815 CATTGTGCGCGGACATCT 835

RESULT 9
CR382121_09/c
WPCOMMENT
Sequence split into 11 fragments LOCUS CR382121 Accession CR382121
Fragment Name Begin End
CR382121_00 1 110000
CR382121_01 100001 210000
CR382121_02 200001 310000
CR382121_03 300001 410000
CR382121_04 400001 510000
CR382121_05 500001 610000
CR382121_06 600001 710000
CR382121_07 700001 810000
CR382121_08 800001 910000
CR382121_09 900001 1010000
CR382121_10 1000001 1062590
Continuation (10 of 11) of CR382121 from base 900001 (CR382121 Kluveromyces fragilis str)

Query Match 19.1%; Score 155.8; DB 8; Length 110000;
Best Local Similarity 53.4%; Pred. No. 1.1e-28;
Matches 473; Conservative 0; Mismatches 337; Indels 75; Gaps 4;

QY 1 ATGACTGATCAATTCACCAATTTTATGATTCATGAGCCACTTAACATTTGTCAAGAGTCC 60
DB 84309 ATTTACTGATTTGTTGCTCTTTTATGATTTGATGAGACGTAACCATGTCATGCTGCT 84250
QY 61 TGTGTGTTTATGCTGTAAGCTTTTATCAAGGTTTGTGGCTTACGCTTGTGACATTTGCT 120
DB 84249 GCGGTGTGTTGCTGCACTTTGTGCAAGGCAATTTGTGCAAGGGAATCAAGACTAGCA 84190
QY 121 TTGCTGATATCAACCAAGAAAGATGCTGCAAGGCAATTTGATCAAGAAATATGCT 180
DB 84189 TTGATGATCTTTCTTGTGCTTATGAGAAATTTAAATGTAATTTGAACTCTTC--- 84133

Query Match	11.3%	Score 92.6;	DB 8;	Length 110000;
Best Local Similarity	52.5%;	Pred. No. 1.1e-12;		
Matches 309;	Conservative 0;	Mismatches 259;	Indels 21;	Gaps 4;

Oy		182	CTGAAAGATTAAGAGATTGTTTCAACGAGTGGAATTCAGGGCATTACAAGTTAATAATTTCCG	92876
Db		92817	CTGAAATACTTAAACCAAGACCTTCGGAGTGAATCCAAGGCATTACAAGTTAATAATTTCCG	92876
Oy		242	ATTCTGATACCCTTCACAAGCGTTTGGCTCAAGTGTCTAAGATTGTTGGTAGTTGCCAT	301
Db		92877	ACCCAAGAATGTGAAAAAGTGATCGACAATAATTGAGAAGATTTCGGCACTATTACG	92936
Oy		302	TGCATTGTGTTAACACAGCTGGTTACTGTGAAAACCTTCCCATG---TGAAGATTACCAG	358
Db		92937	TTTTTGTCCGTATATGCTGTGTGTGCCATGACCGAAGGCAGAAATTTAAGCGAGTGGGT	92996
Oy		359	CCAAGAAGCGTGAAGAGATGTGAAGGTAACTGTGGGTTCTTGTATGTTTCTCAAG	418
Db		92997	ACGATGCGTGGAAAAAGATTGTAGATTGGACTTGTAGTGAAGTTTACTACGCGCTAAGA	93056
Oy		419	CCTTGTCTAAGCCATTGATCAAAGAGTATCAAGGGTGCCTTCTGTGTTTGAATTGGTT	478
Db		93057	CTGTGCGCAAGATTTTCAAGACGCGAA-----AGGGTTCCTGTGTGTTAACAGCAT	93110
Oy		479	CTATGTCTGTGGCATTTGTCCACAGATCTCCAAACCAAGTTGTCTACAACATGTCCAAG	538
Db		93111	CAATGTCTGGCACATTGTGAATGTTCACAAATGTCAAGTCAATCAACAGCTGTAAAG	93170
Oy		539	CTGGTGTATTACCAATTTGGCTAAGACTTTGGCTGTGTGAATGGGCTAAGTACACATCAAG	598
Db		93171	CGGGTGTATTGCACATTGGGTAAGTCTTATGCTATGAATGGGCTCATTCGCT--AGAG	93227
Oy		599	TTAATTTCTTAAACCAAGGTTACATCAAGGCTTTTGACCAAAGATGTATCAATAGTA	658
Db		93228	TTAACACATTTCTCCAGGTTTAAT-----TGTCAAGAGATTTCOGATTTGTTC	93278
Oy		659	ACGAAGAAATTGTCAACAGATGTATCTGTGATTCACAAACAAGATGTCCGAACCAA	718
Db		93279	CTGCGGAAGTCAGGCTTAATATGTGGCAATTGATCCATTATAGAAAGAGAAGCACTTGCTC	93338
Oy		719	AGGAATACATTTGGTGTCTGTTTGTACTGTCTTCTGAATCTGTGCTTCAATCACTACTG	778
Db		93339	AAGAGTTGTGTTGGTGCTTACTTATATTTTGATCCGA--TGCTTACTTACACGACTG	93395
Oy		779	GTCGCAAGTTTACTGTGTGATGTGTGTTTCAAC	809
Db		93396	GCTCGACTTATTTGTGATGTGTGTTTACTC	93426
RESULT 13				
AF002134				
LOCUS AF002134 4293 bp DNA linear PLN 26-JUN-1998				
DEFINITION Candida albicans Souzp (SOU2), Soulp (SOU1) and Vma8p (VMA8) genes,				
complete cds.				
ACCESSION AF002134				
VERSION AF002134.1 GI:2183241				
KEYWORDS				
SOURCE				
ORGANISM				
Candida albicans				
Candida albicans				
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;				
Saccharomycetales; mitosporic Saccharomycetales; Candida.				

[illegible]

REFERENCE 1 (bases 1 to 4293)
AUTHORS Jandon,G., Sherman,F. and Ruschenko,E.
TITLE Monosomy of a specific chromosome determines L-sorbose utilization:
a novel regulatory mechanism in *Candida albicans*
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (9), 5150-5155 (1998)
MEDLINE 98226783
PUBMED 9560244
2 (bases 1 to 4293)
AUTHORS Jandon,G., Ruschenko,E. and Sherman,F.
TITLE Direct Submission
JOURNAL Submitted (02-MAY-1997) Department of Biochemistry and Biophysics,
University of Rochester, 601 Elmwood avenue P.O. Box 712,
Rochester, NY 14642, USA
FEATURES
SOURCE Location/Qualifiers
1. 4293
/organism="Candida albicans"
/mol_type="genomic DNA"
/strain="Sor17"
/db_xref="taxon:5476"
/chromosome="4"
356. 1199
/gene="SOU2"
356. 1198
/gene="SOU2"
/feature="short-chain alcohol dehydrogenase homolog"
/codon_start=1
/transl_table=12
/product="Soulp"
/protein_id="AAC24462.1"
/db_xref="GI:2183242"
/translation="MSKETTSYTNKAGLPPTKATTPDNLDAFSLKGVKASVTGSS
GGIGMAVAEGYAGQADVAIWNHSPADDAEYLTKYGVSKAYKCNVTDPOVERK
VKQIESPFGTIDIFVANAGVAMTEGPEKFDVGVKMNKVVNDLNSVYCAHVGP
RKQKSGVIFPTASMSASIVNVPOLAAVNAKAGVKKLSKFTVEMAPFARVNSVSP
YIATHLEBPADPVKSKWLGTLPIGREKPRELIVGAYLYLASDASTTGADLAVDGS
YTVI"
2063. 2909
/gene="SOU1"
2063. 2908
/gene="SOU1"
/feature="short-chain alcohol dehydrogenase homolog"
/codon_start=1
/transl_table=12
/product="Soulp"
/protein_id="AAC24463.1"
/db_xref="GI:2183243"
/translation="MSREIISFTNPALGPLPTKAPPSNVLDLSLKGVKASVTGSS
GGIGMAVAEAPQAGAVAIWNSKPADDAEYLTEKGVKAKYKCNVTDPDVSKY
INLEKDFGTIDIFVANAGVAMTDPFIDVQYDQMKIVDCDNGVYCAHVGP
KKNKSGSLITTSMSGTIVNIPOLQAPYNAKAKCTHAKSLSYEMAFARVNSISP
GYILTDIADPADPEMKKMWQLTPIGREGLPQELVGAIVLYLASASTTGSNIAVDG
GYTCP"
complement (3073. 3877)
/gene="VMA8"
complement (3074. 3877)
/gene="VMA8"
/feature="Vacuolar H+-ATPase subunit 8 homolog"
/codon_start=1
/transl_table=12
/product="Vma8p"
/protein_id="PAC24464.1"
/db_xref="GI:2183244"
/translation="MSGANREQVFPPTMTLVGVKSKLKAQOQSHLLKKREALT
FRDITRIDDAKRGVMOTAFSLAIVQATGDNISYQVQSVKARFVAKORN
VSGVFLTPDSHINEDVNDPRLTALRAGGQGVOKAKLISKAVETTLASLPATFI
LDVRIKTRNRVNAIEHVIIIPRTNTAIYINGELDEKDRSEPRFLKVOEKQENANA
ABQEBALAKAKBAGATDELAIQDQVLELDIKADEEDVILQEKEDVIF"

gene
CDS
10.7%; Score 87.2; DB 8; Length 4293;
Best Local Similarity 49.7%; Pred. No. 2.9e-11;
Matches 314; Conservative 0; Mismatches 303; Indels 15; Gaps 3;

ORIGIN

QY 182 CTGAAGATTGAAGTTGAAAAGATTCCAAAGATGGCTTCAATAGCTGTGATATTTCG 241
Db 2286 CTGAATATTTAATCGAATAAATAATGATGTCAAAGCCAAAGCTTAAATGATATGTAACG 2345
QY 242 ATTCTGATACGGTTCACAAGGTTTGTCTCAAGTTGCTTAAGATTGTTGTAAGTCCAT 301
Db 2346 ATCTGAATGATGTTCTTAAAGATGATTAATGAAATGAAAAAGATTTCGTCATATTGTA 2405
QY 302 TGCACTTGTTTAAACACAGCTGTTTACTGTGAATACTCCATGTCGAATACCAAGCA 361
Db 2406 TATTGTTGTCATAGCTGAGATTGATGATGATGATGATGATGATGATGATGATGATGATG 2465
QY 362 AGAAGCTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 421
Db 2466 ATGATCAATGAGAAAAAGATCGTTGATGATGATGATGATGATGATGATGATGATGATG 2522
QY 422 TTGTAAGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 481
Db 2523 ATACCGTGGGCAAAATCTTTAAAAAGATTAATCTGTTCAATTAATTAATTAATTAATTA 2582
QY 482 TGTCGTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTC 541
Db 2583 TGTCAGGACATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTC 2642
QY 542 GTGTTATCCATTGGCTTGAAGCTTGGCTTGTGATGATGATGATGATGATGATGATGATG 601
Db 2643 CATGTAATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2702
QY 602 ATTCTTTAAACCGGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 661
Db 2703 ATTCAATTTCTCCAGGCTTATTA-----TTGACTGATATGCTGATTTTTCCTGATC 2753
QY 662 AAGATTGTCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 721
Db 2754 CAGAAATGAAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2813
QY 722 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 781
Db 2814 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2870
QY 782 CCAGCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 813
Db 2871 CAATATGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2902

RESULT 14
AC105434/c 126105 bp DNA linear HTG 20-MAR-2004
LOCUS Magnaporthe grisea chromosome 7 clone 18L14, *** SEQUENCING IN
DEFINITION PROGRESS ***, 2 ordered pieces.
ACCESSION AC105434
VERSION AC105434.2 GI:45597494
KEYWORDS HTG; HTGS PHASE2.
SOURCE Magnaporthe grisea (anamorph: Pyricularia grisea)
ORGANISM Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Magnaporthe; Magnaportheaceae; Magnaporthe.
REFERENCE 1 (bases 1 to 126105)
AUTHORS Thon,M.R., Mitchell,T., Brown,D.E., Diener,S., Taro,A., Pan,H. and
Dean,R.A.
TITLE The sequence of Magnaporthe grisea chromosome 7
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 126105)
AUTHORS Dean,R.A., Dr., Mitchell,T., Thon,M. Dr and Brown,D.E.
TITLE Direct Submission
JOURNAL Submitted (05-JAN-2002) Plant Pathology - Fungal Genomics
Laboratory, North Carolina State University, 840 Main Campus Drive,
Raleigh, NC 27606, USA
3 (bases 1 to 126105)
REFERENCE Thon,M.R., Mitchell,T., Brown,D.E., Diener,S., Taro,A., Pan,H. and
Dean,R.A.
TITLE Direct Submission
JOURNAL Submitted (20-MAR-2004) Center for Integrated Fungal Research,

PT producing fermentation products or in conversion of cheap biomass to useful products.

PS Claim 6; SEQ ID NO 1; 15pp; English.

CC The present invention relates to a MDH dependent L-Xylose reductase
CC enzyme protein and its encoding polynucleotide. The invention is useful
CC in producing fermentation products or in conversion of cheap biomass to
CC useful products. The present sequence is an Ambrosiozima monospora NRRL
CC -1484 MDH dependent L-Xylose reductase cDNA, *AX1* (EC 1.1.1.10).

SQ Sequence 816 BP; 217 A; 159 C; 182 G; 258 T; 0 U; 0 Other;

Query Match	Score	DB	Length
100.0%	816	12	816

Best Local Similarity 100.0%; Pred. No. 2.7e-27;
Matches 816; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATACGTGACATCACTTC	CAACTTTT	AGATTTCG	ATCGATGCG	CACTTAAC	CAATGTCAC	AGGTGCC	60
Db	1	ATGACGTGACATCACTTC	CAACTTTT	AGATTTCG	ATCGATGCG	CACTTAAC	CAATGTCAC	AGGTGCC	60
Qy	61	TGTGTGTGTTT	AGCTTAAG	CTTTAATCA	AGGTTT	TGGCTT	AGGTTT	CTGACATTTGCT	120
Db	61	TGTGTGTGTTT	AGCTTAAG	CTTTAATCA	AGGTTT	TGGCTT	AGGTTT	CTGACATTTGCT	120
Qy	121	TTTGCTGATAT	TGACCAAG	AAAAG	CTGTGTC	CAAC	AAAGCG	AATATAC	180
Db	121	TTTGCTGATAT	TGACCAAG	AAAAG	CTGTGTC	CAAC	AAAGCG	AATATAC	180
Qy	181	ACTGAAG	AATTTGA	AGTTTGA	AGAGAGTT	CCAAAG	ATGSGTT	CAATATGCTGATATTTCT	240
Db	181	ACTGAAG	AATTTGA	AGTTTGA	AGAGAGTT	CCAAAG	ATGSGTT	CAATATGCTGATATTTCT	240
Qy	241	GATTTCTGAT	ATCCGTT	CAACAG	GTGTTT	GCTCAAG	TGCTTAAG	AGATTTTGTATATTTGCA	300
Db	241	GATTTCTGAT	ATCCGTT	CAACAG	GTGTTT	GCTCAAG	TGCTTAAG	AGATTTTGTATATTTGCA	300
Qy	301	TTTGCACTT	TGATTAACA	CAGCTG	TTTACT	GTGAAA	CTTCCAT	GTGAGATTTTACCCAGCC	360
Db	301	TTTGCACTT	TGATTAACA	CAGCTG	TTTACT	GTGAAA	CTTCCAT	GTGAGATTTTACCCAGCC	360
Qy	361	AAGAACG	CTGAGAA	AGATG	TGTGA	AGGTTAA	CTTGT	TGGGTTCTTTTGTATGTTTCTCAAGCC	420
Db	361	AAGAACG	CTGAGAA	AGATG	TGTGA	AGGTTAA	CTTGT	TGGGTTCTTTTGTATGTTTCTCAAGCC	420
Qy	421	TTTTGCT	TAAGCATTTG	ATCAAG	AAAGTAT	CAACAG	GGTCTT	CTGTGTTT	480
Db	421	TTTTGCT	TAAGCATTTG	ATCAAG	AAAGTAT	CAACAG	GGTCTT	CTGTGTTT	480
Qy	481	ATGTCCT	GTGTC	CAATTTG	CAACG	ATCTCA	AAAAC	CAAGTTTGTATCAACATGTCCAGGCT	540
Db	481	ATGTCCT	GTGTC	CAATTTG	CAACG	ATCTCA	AAAAC	CAAGTTTGTATCAACATGTCCAGGCT	540
Qy	541	GGTGT	TATTC	CAATTTG	CTAAG	ACTTTT	GGCTT	GTGAATGGGCTTAAGTACACATCAGATT	600
Db	541	GGTGT	TATTC	CAATTTG	CTAAG	ACTTTT	GGCTT	GTGAATGGGCTTAAGTACACATCAGATT	600
Qy	601	AATTCCT	TAAACCC	AGGTTA	CACTAC	TGAGGTC	CTTGA	CCAAABAAGTTATCAATAGGTATAC	660
Db	601	AATTCCT	TAAACCC	AGGTTA	CACTAC	TGAGGTC	CTTGA	CCAAABAAGTTATCAATAGGTATAC	660
Qy	661	GAAG	AATTTGT	TACAAC	AGATG	ATCTGT	GTATCC	CAACAAGAGATGTCCGAACCAAG	720
Db	661	GAAG	AATTTGT	TACAAC	AGATG	ATCTGT	GTATCC	CAACAAGAGATGTCCGAACCAAG	720
Qy	721	GAAT	ACAT	TG	TGTGCT	TTTGT	ATCTTG	CTTGAATCTGTCTTCAATCACTATCTGT	780
Db	721	GAAT	ACAT	TG	TGTGCT	TTTGT	ATCTTG	CTTGAATCTGTCTTCAATCACTATCTGT	780
Qy	781	GCC	AGCT	TAC	TG	GTG	TATG	GTGATTTCACTTTCTGG	846
Db	781	GCC	AGCT	TAC	TG	GTG	TATG	GTGATTTCACTTTCTGG	846

RESULT 2

ID AAF07518 standard; cDNA; 3203 BP

AC AAF07518;

DT 13-MAR-2001 (first entry)

DE Fusarium venenatum EST SEQ ID NO:41.

KW Multiple gene expression; filamentous fungal cell; EST;

KW expressed sequence tag; fusarium venenatum; Aspergillus niger;
KW Aspergillus oryzae; Trichoderma reesei; identification; recombinant;

KW metabolic pathway e

OS *Fusarium venenatum*

PN WO2000056762-
XX

PD 28-SEP-2000.

PF 22-MAR-2000; 2000WO-US007781.
XX

PR 22-MAR-1999; 99US-00273623.

PA	(NOVO)	NOVO NORDISK BIC
PA	(NOVO)	NOVO NORDISK AS

[illegible]

XX
XX
WBT: 2000-594572/56

XX
XX
Monitoring differential expression of genes in filamentous fungal cells

PT uses fluorescence-labeled nucleic acid substrate of expressed hemance tag

XX
PS CJajm 86: Page 383-384: 3161m: Encji

XX
XX
The present invention describes a method

expression of genes in a first filamentous fungal (FF) cell relative to expression of the same genes in one or more second filamentous fungal cells. The method uses fluorescence-labeled nucleic acids isolated from the FF cells and a substrate of expressed sequence tags (EST). The ESTs are used in the methods for monitoring differential expression of genes in a first filamentous fungal (FF) cell relative to expression of the same genes in one or more second filamentous fungal cells. Monitoring the global expression of genes from FF cells allows the production potential of the microorganisms to be improved. New genes may be discovered, and possible functions of unknown open reading frames can be identified and gene copy number variation and stability can be monitored. The expression of genes can be used to study how FF cells adapt to changes in culture conditions, environmental stress, spore morphogenesis, recombination, metabolic or catabolic pathway engineering. Using ESTs provides several advantages over genomic or random cDNA clones including elimination of redundancy as one spot on an array equals one gene or open reading frame, and organisation of the microarrays based on function of the gene and products to facilitate analysis of the results. AAF07478 to AAF11247 represents ESTs from *Fusarium venenatum*; AAF11248 to AAF11853 represents ESTs from *Aspergillus niger*; AAF11854 to AAF14878 represents ESTs from *Aspergillus oryzae*; and AAF14879 to AAF15337 represents ESTs from *Trichoderma reesei*, which are all specifically claimed in the present invention.

SQ Sequence 3203 BP; 770 A; 861 C; 750 G; 815 T; 0 U; 7 Other;

Query Match	12.2%	Score 99.4	DB 3	Length 3203
-------------	-------	------------	------	-------------

Best Local Similarity 52.7%; Pred. No. 4,5e-18;
Matches 265; Conservative 0; Mismatches 231; Indels 7; Gaps 2

308 TGGTTAACACAGCTGGTTACTGTGAAACTTCCCATGTGAGATTACCACGCGAAGACG 367

Db 2614 TAGTACGCGGCTGGCTTCAAGAGAAATTGAGAGCATCACTATCCATCATGCTGA 2673
 Qy 368 CTGAGAAAGATGTTGAAGTTAACTTTGGGCTTTCTGTATGTTTCTCAAGCTTTGCTA 427
 Db 2674 TCGCGAAGTGTGGGGTGTAAATGTTGACGGTACTCTCTTTCGAGTTGACGTTGCCA 2733
 Qy 428 AGCCATGATCAAAAGAGTATCAAGGGCTTCTGTTGTTGATGTTGTTCTATGCTG 487
 Db 2734 A---ACATCTCAATGAAAGCCAGGATCCTGTGATGATCTGTGTTATGGAACATGCTG 2790
 Qy 488 GTGCCATTGTCAACGATCTCTCAAAACCAAGTTGTCTACAAACATGTCGAAGCTGGTGA 547
 Db 2791 GTGCTATGTCTCAATGTCTCAACGACCAAGCAACATTAACCGGCGCAAGGAGCTGTT 2850
 Qy 548 TCCATTGGCTAAGACTTGTGGCTGTGATGAGTGGCTAAGTCAACATCAGATTTCTT 607
 Db 2851 GTCACTGTGCTGCTTCCCTCCGATGAGTGGGCTCAAGCTGGAATCCGATCACTGTA 2910
 Qy 608 TAAACCCAGGTTACATCAAGCTCTTTCACCAAGATTTTCAATGTTACGAAAGAT 667
 Db 2911 TCTCTCTGCTCAATGTTGACTGCTTAAACGAGAAATTTCTGAACATTAACCCAGATA 2970
 Qy 668 TGTACAAAGATGATCTCTGTGATCCCAACAAGAAATGTCGAAACCAAGAAATCA 727
 Db 2971 TCGAAGAGACCTGGACATCTCTTATCTCAAGGAGCTATGAGACTTACTCAAGATCTGA 3030
 Qy 728 TTGCTGCTGTTTGTACTGCTTCTGATCTGCTGCTTCACTAATCTGCTGCTGCT 787
 Db 3031 TGGGCTCTGT---GACCTTGTGTTGAGATGCGCTTCTTAATGATGAGGCGAGATC 3086
 Qy 788 TACTGCTGATGATGCTGTTTCACT 810
 Db 3087 TTAGAGTAGATGAGAGATATACT 3109

 RESULT 3
 AAF12409
 ID AAF12409 standard; cDNA; 687 BP.
 AC AAF12409;
 XX
 DT 13-MAR-2001 (first entry)
 DE Aspergillus oryzae EST SEQ ID NO:4932.
 XX
 KM Multiple gene expression; filamentous fungal cell; EST;
 KM expressed sequence tag; Fusarium venenatum; Aspergillus niger;
 KM Aspergillus oryzae; Trichoderma reesei; identification; recombination;
 KM culture condition; environmental stress; spore morphogenesis;
 KM metabolic pathway engineering; catabolic pathway engineering; ss.
 OS Aspergillus oryzae.
 XX
 PN WO200056762-A2.
 PD 28-SEP-2000.
 XX
 PP 22-MAR-2000; 2000WO-US007781.
 XX
 PR 22-MAR-1999; 99US-00273623.
 XX
 PA (NOVO) NOVO NORDISK BIOTECH INC.
 PA (NOVO) NOVO NORDISK AS.
 XX
 PI Beika RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;
 DR WPL; 2000-594572/56.
 PT Monitoring differential expression of genes in filamentous fungal cells
 PT uses fluorescence-labeled nucleic acids isolated from the cells and a
 PT substrate of expressed sequence tags.
 XX
 PS Claim 88; Page 2078; 3161pp; English.

XX
 CC The present invention describes a method for monitoring differential
 CC expression of genes in a first filamentous fungal (FF) cell relative to
 CC expression of the same genes in one or more second filamentous fungal
 CC cells. The method uses fluorescence-labeled nucleic acids isolated from
 CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs
 CC are used in the methods for monitoring differential expression of genes
 CC in a first filamentous fungal (FF) cell relative to expression of the
 CC same genes in one or more second filamentous fungal cells. Monitoring the
 CC global expression of genes from FF cells allows the production potential
 CC of the microorganisms to be improved. New genes may be discovered,
 CC possible functions of unknown open reading frames can be identified and
 CC gene copy number variation and stability can be monitored. The expression
 CC of genes can be used to study how FF cells adapt to changes in culture
 CC conditions, environmental stress, spore morphogenesis, recombination,
 CC metabolic or catabolic pathway engineering. Using ESTs provides several
 CC advantages over genomic or random cDNA clones including elimination of
 CC redundancy as one spot on an array equals one gene or open reading frame,
 CC and organization of the microarrays based on function of the gene
 CC products to facilitate analysis of the results. AAF07478 to AAF11247
 CC represents ESTs from Fusarium venenatum; AAF11248 to AAF11853 represents
 CC ESTs from Aspergillus niger; AAF11854 to AAF14878 represents ESTs from
 CC Trichoderma reesei, which are all specifically claimed in the present
 CC invention
 XX
 SQ Sequence 687 BP; 152 A; 210 C; 175 G; 147 T; 0 U; 3 Other;

 Query Match 10.9%; Score 88.6; DB 3; Length 687;
 Best Local Similarity 49.8%; Pred. No. 3.3e-15;
 Matches 279; Conservative 0; Mismatches 275; Indels 6; Gaps 2;

 Qy 196 TTGAAAGAGTTCCAAAGATGGGTTCAATATCTGTGATATTTCTGATTTGATACCGTT 255
 Db 114 TTGAAACAAATGCCCAAGTCAACGCCCACTACGCTGATGTTTCCGACCTTACCTCCGTC 173
 Qy 256 CACAAGGTGTTGCTCAAGTTGCTAAGATTTGGTAAAGTTCACCTGCACTGGTTAAAC 315
 Db 174 AACGATGCCCTCCGATATATATCTCAAGCAAGGCAAGAT---GACACATCGGTCAACC 230
 Qy 316 ACAGTGTGTTACTGTGAAATATTTCCCATGTAAATTTCCAGCCAGAACGCTGAGAAAG 375
 Db 231 TCCGCCGATTCACGAAACCTTGATGCACTCTCACTCAAGCCGCTGCTCAAAAG 290
 Qy 376 ATGGGAAGTTAACTTGGGTTCTTGTGATGTTTCTCAAGCCTTGTCAAGCAATG 435
 Db 291 CTTTGGGCGCTTAATGTCATGATGAAACATCTTTTCCGACCGGTGCGCAAGC---AC 347
 Qy 436 ATCAAGAGATATCAAGGGTCTTCTGTTGTTGATGTTGTTCTATGCTGTGCCATT 495
 Db 348 CTCATGAGGCGCAAGTTCCGGGACGATGTCATGATGTTGATGATGTCGTGCTATC 407
 Qy 496 GTCAACGATCTCTCAAAACCAAGTTGTCTCAACATGTCGAAGGCTGTGTTATTCATTG 555
 Db 408 GACACGCGCCGACGCCAGGCTCTTACCAACCGCCGCAAGGCGCGTGTCTCAACTT 467
 Qy 556 GCTAAGACTTTGGCTTGAATGGCTAAGTACCAATCAAGTGAATTTCTTAAACCA 615
 Db 468 GCCCGTACTTCCCGCGGAAATGGCGGCTACACATTCGGGAGCTGATAGACCTT 527
 Qy 616 GGTACATCTAAGCTCTTTCGACCAAGATGTTATCAATGTTACGAAATTTATCAAC 675
 Db 528 CGATACATGCTTACTGCTGACCGGACGCAAGTTTGTGATGAAACCCCGCATTCGGGAC 587
 Qy 676 AGATGATCTGTGATCCCAACAAGATGTCGAAACCAAGAAATATACATGTTGCT 735
 Db 588 AAGTGATCTGCTCATTCGCCACCGGCAAGATGTTACTCCGAGGACCTGATGGTCCC 647
 Qy 736 GTTTGTACTGCTTCTGTA 755
 Db 648 GTTACCTTTGCTCAAGGA 667

XX	RESULT 4
AD	AD861049
ID	AD861049 standard; cDNA; 637 BP.
XX	
XX	AD861049;
DT	02-DEC-2004 (first entry)
XX	
DE	Bacterial polynucleotide #13036.
XX	
KW	Recombinant DNA construct; improved plant property;
KW	cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW	pathogen tolerance; pest tolerance; plant disease resistance;
KW	cell cycle pathway modification; plant growth regulator;
KW	homologous recombination; seed oil yield; protein yield; carbohydrate;
KW	nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX	bacterial polynucleotide; gene; ss.
XX	
OS	Bacteria.
XX	
PN	US2003233675-A1.
XX	
PD	18-DEC-2003.
XX	
PF	20-FEB-2003; 2003US-00359493.
XX	
PR	21-FEB-2002; 2002US-0360039P.
XX	
PA	(CAOY/) CAO Y.
PA	(HINK/) HINKLE G J.
PA	(SLAT/) SLATER S C.
XX	(CHEN/) CHEN X.
PA	(GOLD/) GOLDMAN B S.
XX	
PI	Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
DR	WPI; 2004-061375/06.
XX	
PT	New recombinant DNA construct comprising a promoter positioned to provide
PT	for expression of a polynucleotide encoding a polypeptide from a
PT	microbial source, useful for producing plants with improved properties.
XX	
ES	Claim 1; SEQ ID NO 36723; 122pp; English.
XX	
CC	The invention relates to a recombinant DNA construct comprising a
CC	promoter functional in a plant cell, where the promoter is positioned to
CC	provide for expression of a polynucleotide encoding a polypeptide from a
CC	microbial source. The invention also relates to a transformed plant
CC	comprising the recombinant DNA construct and a method of producing a
CC	transformed plant having an improved property. The plant is a crop plant
CC	such as maize or soybean. The method of producing a transformed plant
CC	having an improved property comprises transforming a plant with the
CC	recombinant DNA construct and growing the transformed plant, where the
CC	polynucleotide or polypeptide is useful for improving plant properties.
CC	The recombinant DNA construct is useful for producing plants with
CC	improved plant properties, e.g. improved cold, heat or drought tolerance,
CC	tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC	increased resistance to plant disease, better growth rate by modification
CC	of the cell cycle pathway with plant growth regulators, increased rate of
CC	homologous recombination, modified seed oil or protein yield and/or
CC	content, improved yield by modification of carbohydrate, nitrogen or
CC	phosphorus use and/or uptake, by modification of photosynthesis or by
CC	providing improved plant growth and development under at least one stress
CC	condition, improved lignin production or improved galactomannan
CC	production. This sequence represents a bacterial polynucleotide used in
CC	the scope of the invention. Note: The sequence data for this patent did
CC	not form part of the printed specification but was obtained in electronic
CC	format from USPTO at seqdata.uspto.gov/sequence.html .
XX	
SQ	Sequence 637 BP; 149 A; 177 C; 169 G; 142 T; 0 U; 0 Other;
XX	
Query Match	10.0%; Score 81.8; DB 13; Length 637;
Best Local Similarity	55.0%; Pred. No. 3.1e-13;

Matches	183;	Conservative	0;	Mismatches	147;	Indels	3;	Gaps	1;
QY	304	CAC TTG GTT AAC	CAC AG CTG TTA CTG TAA	AAC TTT CCA TGT AAG	ATT ACC CAC GGC	CAA G	363		
Db	70	CAT CTG CTA CGT CTG	CGG GAT TTA ACC GAA	AA TCT CGA CGC	CAT CAG CTAC CCA TAC	CA C	129		
QY	364	AA CCG TGA AAG	ATG GTT AAC GTT AAC	CTG TTT GGG ATT CTT	GTG TAT GAT TTT CCA GGC	CTT	423		
Db	130	CG TAT GCA GAA GCT	CTG GGG GGT AAT CCG	ACA CCG CAC AAT ATC	GTG TTG CTA CGG CCG	TG	189		
QY	424	GCT AAG CCA TTG	ATC AAG AAG ATG ATC	CA AAG GGT CTT CTG TTT	GAT TTG ATT GAT TGT CTA TG	483			
Db	190	GCA AAG C	CTT GAT --- GGA GCG	CAG GCG CCG CCG GAT	ATT GTG ATG ATG GGT AAG	CA TG	246		
QY	484	TCT GGT GCA ATG	TCA ACG ATC CTT CAA	AAC CA A GTT GCT TCA	ACA CAT GTT CCA A GCG	TGT	543		
Db	247	TCT GGT GCA ATG	TCA ACG ATC CTT CAA	ACA CCA A GCG CCG	CTT CAA CGC GCG CAG	CAC	306		
QY	544	GTT ATC CAA TTG	GCT AAG ACT TTT GCG	CTT GTG ATG TGG GCT	TAA GTC AACA CAT CAG	AAT	603		
Db	307	GTT GCG CCA C	CTT GCT TCG CTT TGC C	CTC TCA ATG GGA GGA	CAT GGT ATC AGG GTT	AAC	366		
QY	604	TCT TTA AAC	CGA GTT ATC ATC	AGG GTT CCG TTT G	636				
Db	367	TGC ATTA GCG	CGT GAT TAC TGT CTA	CTG CTT G	399				

RESULT 5
 ADS48561
 ID ADS48561 standard; cDNA, 735 BP.
 XX
 AC ADS48561;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Bacterial polynucleotide #3304.
 XX
 KW Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polynucleotide; gene; ss.
 XX
 OS Bacteria.
 XX
 PN US200323675-A1.
 XX
 PD 18-DEC-2003.
 XX
 PF 20-FEB-2003; 2003US-00369493.
 XX
 PR 21-FEB-2002; 2002US-0360039P.
 XX
 PA (CAOY/) CAO Y.
 XX
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 DR WPI; 2004-061375/06.
 XX
 PT New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX
 PS Claim 1; SEQ ID NO 26991; 122pp; English.
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to

CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transgenic plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

SO Sequence 735 BP; 136 A; 234 C; 183 G; 182 T; 0 U; 0 Other;

Query Match 9.9%; Score 80.6; DB 13; Length 735;
Best Local Similarity 61.8%; Pred. No. 7.4e-13;
Matches 128; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 436 ATCAAGAGGATGATCAAGGCTCTGTGTTGTTGATGTTCTATGTCGGCCATT 495
DB 364 ATGATGGCCGCTGGGAAACCCCGCTCCATATCTGTGCTCCATGTCGGCACATC 423
QY 496 GTCAACGATCTCAAAACCAAGTGTCTCAACATGTCACAGCTGTTATTCATTG 555
DB 424 GTCAACTACCTCAGAGAGAGCTGCTCAACAGCCTCTAAGCGTGGTCAATCACTC 483
QY 556 GCTAAGACTTTGGCTTGTGATGGCTAAGTCAACATCAAGATTATCTTTAAACCA 615
DB 484 GCGAAGTCTCTTGTGCGGAGTGGGCGCAAGTACATCCGTTTACATGCAATTTCCCC 543
QY 616 GGTTCATCTACGAGGTCCTTTGACCAAG 642
DB 544 GGTTCATGATGATACCGCCCTGATTAAG 570

RESULT 6

AAH74598
ID AAH74598 standard; DNA; 852 BP.

AAH74598;

15-OCT-2001 (first entry)

Synthetic nucleotide sequence of a NADPH-dependent carbonyl reductase.

Methionine gamma-lyase; mda gene; free folding energy; gene shuffling;

directed evolution; molecular breeding;

NADPH-dependent carbonyl reductase; ss.

Candida magnoliae.

WO200155342-A2.

02-AUG-2001.

31-JAN-2001; 2001WO-US003186.

31-JAN-2000; 2000US-00494921.

08-DEC-2000; 2000US-00734237.

(BIOC-) BIOCATALYTICS INC.

PI Rozzell DJ, Bui P, Hua L;
XX WPI; 2001-483235/52.

PT Designing synthetic nucleic acid sequences for improved amplification,
PT expression in host cell, by comparing free energy of folding of a
PT starting polynucleotide and a modified polynucleotide having a codon
PT replacement.

PS Claim 9; Page 104; 117p; English.

CC The present sequence encodes a modified NADPH-dependent carbonyl
CC reductase. The polynucleotide was modified using the method of the
CC invention. The specification describes a method for designing a synthetic
CC polynucleotide. The method comprises providing a starting polynucleotide,
CC determining the predicted free energy of folding per base of the
CC polynucleotide, modifying the polynucleotide by replacing a codon with a
CC different codon to provide a modified polynucleotide, determining free
CC energy of folding per base of the modified polynucleotides, and comparing
CC this with that of the original polynucleotide. The method is useful for
CC developing nucleic acid sequences that enhance expression of the encoded
CC protein in a heterologous host. The design and preparation of the
CC synthetic genes are used in application of gene shuffling, directed
CC evolution and molecular breeding methods. The method allows expression of
CC genes from various organisms such as mammals, plants, yeast, fungi and
CC bacteria in prokaryotic hosts, such as *Escherichia coli* and eukaryotic
CC hosts at commercially viable levels, in particular proteins with low
CC yield such as methionine gamma-lyase from *P. putida*

SO Sequence 852 BP; 196 A; 216 C; 205 G; 235 T; 0 U; 0 Other;

Query Match 9.1%; Score 74.6; DB 5; Length 852;
Best Local Similarity 56.0%; Pred. No. 4.4e-11;
Matches 209; Conservative 0; Mismatches 149; Indels 15; Gaps 3;

QY 437 TCAAGAGGATGATCAAGGCTCTGTGTTGTTGATGTTCTATGTCGGCCATTG 496
DB 485 TCGAAAGAGAGTAAAGAGCGCTGTGTTTCAACCGCTTCTATGTCGGTCAATCG 544
QY 497 TCAAGATCTCAAAACCAAGTGTCTCAACATGTCACAGCTGTTATTCATTG 556
DB 545 TTAAGATACCGCAGTTTCAAGCTCAACAGCTGCTAAAGCTGTGTTGTCATTGG 604
QY 557 CTAAAGACTTTGGCTTGTGATGGCTAAGTCAACATCAGATTAACTTTAAACCCAG 616
DB 605 CTAAATCTCTGGCTGTAATTCGCTCGT---TCGCTGCTTAACTCTGTTCTCCGG 661
QY 617 GTTACATCTACGGTCTTTGACCAAGATGTTATCAATGTAACGAAGATTGTAACA 676
DB 662 GCTACATCAAC-----ACGAAATCTCTGACTTTGTACCGAAGAACTGAGAAC 712
QY 677 GATGATCTCTGATATCCCAACAAAGATGTCGAACCAAGATATCTGGTCTG 736
DB 713 AATGATGCTCTGTGATCCGCTGCGCGTGTGACGAAGCTGAACTGTGTTGGTCTT 772
QY 737 TTTGTACTTCTTCTGATCTGCTGCTTCACTACACTAGTGGCCAGTTACTGCTTG 796
DB 773 ACTGTTTCTGCTTCTTGA---CGCTGTTTCTACGCTACCGGCACTGACATCATGTTG 829
QY 797 ATGGTGGTTTCAAC 809
DB 830 ACGGTGGTTACAC 842

RESULT 7

AAF12302
ID AAF12302 standard; cDNA; 1146 BP.

AAF12302;

13-MAR-2001 (first entry)

Aspergillus oryzae EST SEQ ID NO:4825.

XX Multiple gene expression; filamentous fungal cell; EST;
 KM expressed sequence tag; *Fusarium venenatum*; *Aspergillus niger*;
 KM *Aspergillus oryzae*; *Trichoderma reesei*; identification; recombination;
 KM culture condition; environmental stress; spore morphogenesis;
 KM metabolic pathway engineering; catabolic pathway engineering; ss.
 XX *Aspergillus oryzae*.
 OS
 PN NO200056762-A2.
 XX
 XX 28-SEP-2000.
 PD
 XX 22-MAR-2000; 2000WO-US007781.
 PE
 XX 22-MAR-1999; 99US-00273623.
 PR
 XX (NOVO) NOVO NORDISK BIOTECH INC.
 PA (NOVO) NOVO NORDISK AS.
 XX
 PI Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olesen PB;
 XX WPI; 2000-594572/56.
 DR
 XX Monitoring differential expression of genes in filamentous fungal cells
 PT uses fluorescence-labeled nucleic acids isolated from the cells and a
 PT substrate of expressed sequence tags.
 XX
 PS Claim 88, Page 2039, 3161pp; English.
 XX
 XX The present invention describes a method for monitoring differential
 CC expression of genes in a first filamentous fungal (FF) cell relative to
 CC expression of the same genes in one or more second filamentous fungal
 CC cells. The method uses fluorescence-labeled nucleic acids isolated from
 CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs
 CC are used in the methods for monitoring differential expression of genes
 CC in a first filamentous fungal (FF) cell relative to expression of the
 CC same genes in one or more second filamentous fungal cells. Monitoring the
 CC global expression of genes from FF cells allows the production potential
 CC of the microorganisms to be improved. New genes may be discovered,
 CC possible functions of unknown open reading frames can be identified and
 CC gene copy number variation and stability can be monitored. The expression
 CC of genes can be used to study how FF cells adapt to changes in culture
 CC conditions, environmental stress, spore morphogenesis, recombination,
 CC metabolic or catabolic pathway engineering. Using ESTs provides several
 CC advantages over genomic or random cDNA clones including elimination of
 CC redundancy as one spot on an array equals one gene or open reading frame,
 CC and organization of the microarrays based on function of the gene
 CC products to facilitate analysis of the results. AAF07478 to AAF11247
 CC represents ESTs from *Fusarium venenatum*; AAF11248 to AAF11853 represents
 CC ESTs from *Aspergillus niger*; AAF11854 to AAF14878 represents ESTs from
 CC *Aspergillus oryzae*; and AAF14879 to AAF15337 represents ESTs from
 CC *Trichoderma reesei*, which are all specifically claimed in the present
 CC invention.
 CC
 SQ Sequence 1146 BP; 262 A; 313 C; 274 G; 297 T; 0 U; 0 Other;
 XX
 XX
 Query Match 9.1%; Score 74.2; DB 3; Length 1146;
 Best Local Similarity 50.6%; Pred. No. 6, 6e-11;
 Matches 206; Conservative 0; Mismatches 198; Indels 3; Gaps 1;

QY 529 ATGTCCAGGCTGTGTATTCATTGGCTTAAGACTTTGGCTTGTGATGGCTTAAGTAC 588
 Db 744 GTTAGCAGGCTCAGGCTTCACTTGGCCCGTAACTCGCATGATGAGGCGCCGTAC 803
 QY 589 AACATCAGATTAATCTTTAAACCCAGGTTACATCAGGTCCTTTGACCAAGATGTT 648
 Db 804 AACATTCAGTCAACACCACTCTGCCCGGCTTAATTTACTCCATGCTTGAGGAAGCTC 863
 QY 649 ATCAATGTTAAGCAGAGATTGTACAACAGATGATCTTGTATGCCACAAGAAATG 708
 Db 864 TTGTTGAGTTCCCTGAGCGTCGCGAGATGCGCCAAACATTAATGCTGGGACGCTG 923
 QY 709 TCGAACCAAGAAATACATTGCTGCTGTTTGTACTGCTTTCTGA 755
 Db 924 TCTACCCCTTAACGAGTACCGGTGGCGCTGCGCTTCTTCTCAGTGA 970
 RESULT 8
 ADS47380
 ID ADS47380 standard; cDNA; 783 BP.
 AC
 XX ADS47380;
 XX
 XX 02-DEC-2004 (first entry)
 DT
 XX Bacterial polynucleotide #2123.
 DE
 XX Recombinant DNA construct; transformed plant; improved plant property;
 XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KM pathogen tolerance; pest tolerance; plant disease resistance;
 KM cell cycle pathway modification; plant growth regulator;
 KM homologous recombination; seed oil yield; protein yield; carbohydrate;
 KM nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KM bacterial polynucleotide; gene; ss.
 XX
 XX Bacteria.
 OS
 XX US200323675-A1.
 PN
 XX 18-DEC-2003.
 PD
 XX 20-FEB-2003; 2003US-00369493.
 PF
 XX 21-FEB-2002; 2002US-0360039P.
 PR
 XX
 XX (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX WPI; 2004-061375/06.
 DR
 XX New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 PT
 PS Claim 1; SEQ ID NO 25810; 122pp; English.
 XX
 XX The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,


```

PN US200323675-A1.
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
DR WPI, 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 26610; 122bp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomanan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 774 BP; 226 A; 160 C; 229 G; 159 T; 0 U; 0 Other;
XX
Query Match 7.8%; Score 63.6; DB 13; Length 774;
Best Local Similarity 61.4%; Pred. No. 6,9e-08;
Matches 102; Conservative 0; Mismatches 64; Indels 0; Gaps 0
QY 472 ATTGGTCTAATGTCTGTGGCATTTGCACAGATCTCTCAAACAAGTTGTCACACATG 531
DB 436 ATCGCCTCATATGCCGACACATCTGGAACAAACTCAGAAGACAGAGCTTAACACGCT 495
QY 532 TCCAAGGCGTGCTATTTCACATTTGGCTAAGACTTTGGCTTGTGAAAGGGCTAAGTACAC 591
DB 496 TCGAAGGCGGTGTGATTCATCTCACAGATCTCTGCGCGCCGACAGTGGGCCCTTAGCGGA 555
QY 592 ATCAGATTAAATCTTTAAAACCAAGGTTACATCTAGCGTCTTTGA 637
DB 556 ATCAGGTTAACGACATAAGCCCCGATACATAGAAACACTCTCA 601
RESULT 11
ID AAFL1343 standard; cDNA; 636 BP.
XX
AF11343;

```

XX	13-MAR-2001	(first entry)
XX	Aspergillus niger EST SEQ ID NO:1866.	
DE	Aspergillus niger	
XX	Multiple gene expression; filamentous fungal cell, EST;	
KW	expressed sequence tag; Fusarium venenatum; Aspergillus niger;	
KW	Aspergillus oryzae; Trichoderma reesei; identification; recombination;	
KM	culture condition; environmental stress; spore morphogenesis;	
KM	metabolic pathway engineering; catabolic pathway engineering; ss.	
XX	Aspergillus niger.	
OS	Aspergillus niger.	
PN	WO200056762-A2.	
XX	22-MAR-2000; 2000WO-US007781.	
XX	28-SEP-2000.	
XX	22-MAR-1999; 99US-00273623.	
PR	(NOVO) NOVO NORDISK BIOTECH INC.	
PA	(NOVO) NOVO NORDISK AS.	
PI	Berka RM, Rey MW, Shuter JR, Kaupinen S, Clausen IG, Olsen PB;	
PI	WPI, 2000-594572/56.	
DR	Monitoring differential expression of genes in filamentous fungal cells	
PT	uses fluorescence-labeled nucleic acids isolated from the cells and a	
PT	substrate of expressed sequence tags.	
XX	Claim 87; Page 1728; 3161pp; English.	
XX	The present invention describes a method for monitoring differential	
CC	expression of genes in a first filamentous fungal (FF) cell relative to	
CC	expression of the same genes in one or more second filamentous fungal	
CC	cells. The method uses fluorescence-labeled nucleic acids isolated from	
CC	the FF cells and a substrate of expressed sequence tags (EST). The ESTs	
CC	are used in the methods for monitoring differential expression of genes	
CC	in a first filamentous fungal (FF) cell relative to expression of the	
CC	same genes in one or more second filamentous fungal cells. Monitoring the	
CC	global expression of genes from FF cells allows the production potential	
CC	of the microorganisms to be improved. New genes may be discovered,	
CC	and possible functions of unknown open reading frames can be identified and	
CC	gene copy number variation and stability can be monitored. The expression	
CC	of genes can be used to study how FF cells adapt to changes in culture	
CC	conditions, environmental stress, spore morphogenesis, recombination,	
CC	metabolic or catabolic pathway engineering. Using ESTs provides several	
CC	advantages over genomic or random cDNA clones including elimination of	
CC	redundancy as one spot on an array equals one gene or open reading frame,	
CC	and organization of the microarrays based on function of the gene	
CC	products to facilitate analysis of the results. AAF07478 to AAF11247	
CC	represents ESTs from Fusarium venenatum; AAF11248 to AAF11853 represents	
CC	ESTs from Aspergillus niger; AAF11854 to AAF14878 represents ESTs from	
CC	Aspergillus oryzae; and AAF14879 to AAF15337 represents ESTs from	
CC	Trichoderma reesei, which are all specifically claimed in the present	
CC	invention	
XX	Sequence 636 BP; 169 A; 155 C; 161 G; 149 T; 0 U; 2 Other;	
XX	Query Match 7.4%; Score 60.6; DB 3; Length 636;	
XX	Best Local Similarity 51.3%; Pred. No. 4, 7e-07;	
XX	Matches 194; Conservative 0; Mismatches 175; Indels 9; Gaps 2	
QY	211 AAGATGGGTTCATATGCTGTGATATTTGATTTGATTCAGTTCACAGAGTGTGGCT 270	
DB	200 AATCTAGCTGCTTAAAGCTATGTCGGAATCAGAGAGATTTGAGGCGTATCCAG 259	
QY	271 CAAATTGCTAAGATTTGGTAAAGTTCATTCGATTCGATTCGATTCGATTCGATTCG 330	
DB	260 CAAATTGCTTCAAGCTTTGGCAAGCTTGATATATTCGTGTGAAC---TCGGGGGTCAC 316	

PS Claim 1; Page 10-11; 14pp; Japanese.
XX
CC The invention relates to gene encoding D-arabinitol dehydrogenase.
CC isolated from *Bacillus* sp. IKD-5A868 strain. The protein can be expressed
CC by standard recombinant methodologies. D-arabinitol dehydrogenase is used
CC as a clinical diagnosing agent for mycosis. The present sequence
CC represents the coding sequence of the D-arabinitol dehydrogenase
XX
SQ Sequence 774 BP; 155 A; 214 C; 271 G; 134 T; 0 U; 0 Other;
Query Match 7.2%; Score 59; DB 3; Length 774;
Best Local Similarity 50.4%; Pred. No. 1.5e-06;
Matches 171; Conservative 0; Mismatches 165; Indels 3; Gaps 1;
QY 478 TCTATGTCGTGGCCATTGTCACAGATCTTAAACCAAGTTGCTTACACATGCCAAG 537
DB 439 TCCATGTCGGGCGCTGATGTCATATACCGCGACGCGGCGGCTTACATGTCCTCAAG 498
QY 538 GCTGTGTATTCATTGGCTTAAGACTTTGGCTTGTGAATGGGCTAAGTACATCAGA 597
DB 499 GCGGGGCTATTATGCTGACCAAGAGCTGGATCCGAATGGGCGGCGGCTGCGC 558
QY 598 GTTATTTCTTTAAACCAAGTTACATCTACGCTCTTGAACCAAGATGTTATCATNGT 657
DB 559 GTCAACAGATTTGGCGCCGGCTACATGAAGCAAGCTGACGGAACCGTATTTGCGCGCC 618
QY 658 AACGAAGATTGTACAACAGATGATCTCTGTATCCCAACAAGAAATGTCGAACCA 717
DB 619 GAGGCGAGATGATGACAAAGTGGCTGTATGACCCGATGAGGCGCGGCGCTTCC- 677
QY 718 AAGGAATACATTGGGCTGTTTGTACTTGTCTTGAATGCTGCTTCACTACTACT 777
DB 678 --GACGAGCTGGGCGGCGGCTTACCTTGTGAGAGCTTCTCTTGGCCACG 735
QY 778 GATGCCAGCTTACTGTTGATGATGATGTTCACTTCTTGG 816
DB 736 GCGGCGGTGTCAGCATTTGACGGGGGCTATACGATTTGG 774
RESULT 14
AAZ46763
ID AAZ46763 standard; DNA; 2255 BP.
XX
AC AAZ46763;
XX
DT 31-MAR-2000 (first entry)
XX
DE *Bacillus* D-arabinitol dehydrogenase encoding genomic DNA.
XX
KM D-arabinitol dehydrogenase; clinical diagnosis; mycosis; ds.
XX
OS *Bacillus* sp.
XX
XX
XX Key Location/Qualifiers
FH CDS 366..1142
FT /tag= a
FT /product= "D-arabinitol dehydrogenase"
FT /note= "the coding sequence is also given in AAZ46762"
XX
PN JP1332569-A.
XX
PD 07-DEC-1999.
XX
PF 26-MAY-1998; 98JP-00143637.
XX
PR 26-MAY-1998; 98JP-00143637.
XX
XX (IKED-) IKEDA SHOKKEN KK.
PA (NIPK) NIPPON KAYAKU KK.
XX
XX MPI: 2000-091353/08.
DR P-PSDB; AA156815.
XX

PT Arabinitol dehydrogenase gene encoding D-arabinitol dehydrogenase -
PT useful as a clinical diagnosing agent for mycosis.
XX
PS Claim 6; Page 12; 14pp; Japanese.
XX
CC The invention relates to gene encoding D-arabinitol dehydrogenase.
CC isolated from *Bacillus* sp. IKD-5A868 strain. The protein can be expressed
CC by standard recombinant methodologies. D-arabinitol dehydrogenase is used
CC as a clinical diagnosing agent for mycosis. The present sequence
CC represents the DNA encoding the D-arabinitol dehydrogenase
XX
SQ Sequence 2255 BP; 418 A; 623 C; 728 G; 486 T; 0 U; 0 Other;
Query Match 7.2%; Score 59; DB 3; Length 2255;
Best Local Similarity 50.4%; Pred. No. 2.4e-06;
Matches 171; Conservative 0; Mismatches 165; Indels 3; Gaps 1;
QY 478 TCTATGTCGTGGCCATTGTCACAGATCTTAAACCAAGTTGCTTACACATGCCAAG 537
DB 804 TCCATGTCGGGCGCTGATGTCATATACCGCGACGCGGCGGCTTACATGTCCTCAAG 863
QY 538 GCTGTGTATTCATTGGCTTAAGACTTTGGCTTGTGAATGGGCTAAGTACATCAGA 597
DB 864 GCGGGGCTATTATGCTGACCAAGAGCTGGATCCGAATGGGCGGCGGCTGCGC 923
QY 598 GTTATTTCTTTAAACCAAGTTACATCTACGCTCTTGAACCAAGATGTTATCATNGT 657
DB 924 GTCAACAGATTTGGCGCCGGCTACATGAAGCAAGCTGACGGAACCGTATTTGCGCGCC 983
QY 658 AACGAAGATTGTACAACAGATGATCTCTGTATCCCAACAAGAAATGTCGAACCA 717
DB 984 GAGGCGAGATGATGACAAAGTGGCTGTATGACCCGATGAGGCGCGGCGCTTCC- 1042
QY 718 AAGGAATACATTGGGCTGTTTGTACTTGTCTTGAATGCTGCTTCACTACTACT 777
DB 1043 --GACGAGCTGGGCGGCGGCTTACCTTGTGAGAGCTTCTCTTGGCCACG 1100
QY 778 GATGCCAGCTTACTGTTGATGATGATGTTCACTTCTTGG 816
DB 1101 GCGGCGGTGTCAGCATTTGACGGGGGCTATACGATTTGG 1139
RESULT 15
AAC42189
ID AAC42189 standard; DNA; 798 BP.
XX
AC AAC42189;
XX
DT 17-OCT-2000 (first entry)
XX
DE *Arbidiopsis* thaliana DNA fragment SEQ ID NO: 34619.
XX
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS *Arbidiopsis* thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
XX
PR 05-MAR-1999; 99US-0123180P.
XX
PR 09-MAR-1999; 99US-0123548P.
XX
PR 23-MAR-1999; 99US-0125788P.
XX
PR 25-MAR-1999; 99US-0126264P.
XX
PR 29-MAR-1999; 99US-0126785P.
XX
PR 01-APR-1999; 99US-0127462P.
XX
PR 06-APR-1999; 99US-0128234P.
XX
PR 08-APR-1999; 99US-0128714P.
XX

PR	16-APR-1999;	99US-0129845P.
PR	19-APR-1999;	99US-0130077P.
PR	21-APR-1999;	99US-0130449P.
PR	23-APR-1999;	99US-0130510P.
PR	23-APR-1999;	99US-0130891P.
PR	30-APR-1999;	99US-0131449P.
PR	30-APR-1999;	99US-0132048P.
PR	30-APR-1999;	99US-0132407P.
PR	04-MAY-1999;	99US-0132484P.
PR	05-MAY-1999;	99US-0132485P.
PR	06-MAY-1999;	99US-0132486P.
PR	07-MAY-1999;	99US-0132487P.
PR	11-MAY-1999;	99US-0132863P.
PR	11-MAY-1999;	99US-0134256P.
PR	14-MAY-1999;	99US-0134218P.
PR	14-MAY-1999;	99US-0134219P.
PR	14-MAY-1999;	99US-0134370P.
PR	18-MAY-1999;	99US-0134768P.
PR	19-MAY-1999;	99US-0134941P.
PR	20-MAY-1999;	99US-0135124P.
PR	21-MAY-1999;	99US-0135353P.
PR	24-MAY-1999;	99US-0135629P.
PR	25-MAY-1999;	99US-0136021P.
PR	27-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136392P.
PR	01-JUN-1999;	99US-0137222P.
PR	03-JUN-1999;	99US-0137528P.
PR	04-JUN-1999;	99US-0137502P.
PR	07-JUN-1999;	99US-0137724P.
PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	14-JUN-1999;	99US-0138847P.
PR	16-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	17-JUN-1999;	99US-0139453P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	21-JUN-1999;	99US-0139763P.
PR	22-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140822P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143544P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	19-JUL-1999;	99US-0144336P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144881P.
PR	21-JUL-1999;	99US-0145008P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145086P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	23-JUL-1999;	99US-0145152P.
PR	23-JUL-1999;	99US-0145154P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	28-JUL-1999;	99US-0145951P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147003P.
PR	04-AUG-1999;	99US-0147204P.
PR	05-AUG-1999;	99US-0147132P.
PR	05-AUG-1999;	99US-0147132P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2005, 12:10:28 ; Search time 185 Seconds

(without alignment)
7217.307 Million cell updates/sec

Title: US-10-720-018-1

Perfect score: 816
Sequence: 1 atgactgactcattccac.....atgctgttcttcttcttg 816

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 1202764 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	238.2	29.2	696	US-08-336-198C-6	Sequence 6, Appli
2	238.2	29.2	696	US-09-184-965-6	Sequence 6, Appli
3	218.6	26.8	846	US-09-248-796A-29	Sequence 29, Appli
4	86.6	10.6	927	US-09-248-796A-2490	Sequence 2490, Ap
5	74.6	9.1	852	US-09-734-237B-71	Sequence 71, Appli
6	61.8	7.6	873	US-09-248-796A-2654	Sequence 2654, Ap
7	55	6.7	336	US-09-248-796A-2488	Sequence 2488, Ap
8	53	6.5	372	US-09-248-796A-6987	Sequence 6987, Ap
9	53	6.5	774	US-09-134-001C-725	Sequence 725, App
10	51.8	6.3	825	US-09-248-796A-2730	Sequence 2730, Ap
11	51.2	6.3	852	US-09-367-012-2	Sequence 2, Appli
12	51.2	6.3	852	US-09-777-157A-2	Sequence 2, Appli
13	51.2	6.3	852	US-09-734-237B-69	Sequence 69, Appli
14	51.2	6.3	951	US-09-489-039A-5819	Sequence 5819, Ap
15	51.2	6.3	1303	US-09-367-012-8	Sequence 8, Appli
16	51.2	6.3	1303	US-09-777-157A-8	Sequence 8, Appli
17	50.4	6.2	4411529	US-09-103-840A-1	Sequence 1, Appli
18	49.8	6.1	873	US-09-248-796A-2489	Sequence 2489, Ap
19	48.2	5.9	3993	US-09-710-279-3985	Sequence 3985, Ap
20	47.4	5.8	4403765	US-09-103-840A-2	Sequence 2, Appli
21	47	5.8	792	US-09-134-001C-1675	Sequence 1675, Ap
22	45.2	5.5	2774	US-09-363-189B-5	Sequence 5, Appli
23	44.4	5.4	1248	US-10-024-806-7	Sequence 7, Appli
24	43	5.3	879	US-09-468-738A-1	Sequence 1, Appli
25	43	5.3	879	US-09-940-019-1	Sequence 1, Appli
26	43	5.3	879	US-09-940-037A-1	Sequence 1, Appli
27	43	5.3	891	US-09-468-738A-22	Sequence 22, Appli

28	43	5.3	891	US-09-940-019-22	Sequence 22, Appli
29	43	5.3	891	US-09-940-037A-22	Sequence 22, Appli
30	43	5.3	1787	US-09-468-738A-17	Sequence 17, Appli
31	43	5.3	1787	US-09-940-019-17	Sequence 17, Appli
32	43	5.3	1787	US-09-940-037A-17	Sequence 17, Appli
33	42.2	5.2	789	US-09-710-279-3157	Sequence 3157, Ap
34	42.2	5.2	3210	US-09-710-279-3922	Sequence 3922, Ap
35	42.2	5.2	3368	US-09-710-279-3922	Sequence 4169, Ap
36	41.4	5.1	825	US-09-134-001C-1594	Sequence 1594, Ap
37	41.4	5.1	916	US-09-710-279-381	Sequence 381, App
38	41.4	5.1	3024	US-09-710-279-3426	Sequence 3426, Ap
39	41.2	5.0	2472	US-09-248-796A-148	Sequence 148, App
40	41	5.0	437	US-09-468-738A-14	Sequence 14, Appli
41	41	5.0	437	US-09-940-019-14	Sequence 14, Appli
42	41	5.0	437	US-09-940-037A-14	Sequence 14, Appli
43	41	5.0	837	US-09-328-352-2671	Sequence 2671, A
44	40.8	5.0	601	US-09-949-016-62984	Sequence 62984, A
45	40.8	5.0	50263	US-09-949-016-13563	Sequence 13563, A

ALIGNMENTS

RESULT 1
US-08-336-198C-6
; Sequence 6, Application US/08336198C
; Patent No. 5866382
; GENERAL INFORMATION:
; APPLICANT: Hallborn, Johan
; APPLICANT: Penttila, Merja
; APPLICANT: Ojamo, Heikki
; APPLICANT: Keranen, Sirka
; APPLICANT: Hahn-Hagedorn, Barbel
; APPLICANT: Waldfriedson, Mats
; APPLICANT: Aikarsinen, Ulla
; TITLE OF INVENTION: XYLOSE UTILIZATION BY RECOMBINANT YEASTS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 301 N. Washington St.
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22046-3487
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,198C
; FILING DATE: 03-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 696 base pairs
; TYPE: nucleic acid
; STRANDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Pichia stipitis
; STRAIN: CBS-6054
; FEATURE:
; NAME/KEY: CDS

```

: LOCATION: 1..693
: OTHER INFORMATION: /standard_name="xy11c1"
: OTHER INFORMATION: dehydrogenase"
: PUBLICATION INFORMATION:
: DOCUMENT NUMBER: FI 901771
: FILING DATE: 06-APR-1990
:
: JS-08-336-198C-6

```

Query Match	29.2%	Score 238.2	DB 2	Length 696
Best Local Similarity	64.4%	Pred. No. 9.4e+65		
Matches 373	Conservative 0	Mismatches 203	Indels 3	Gaps 1

OY	227	CTGTGATTAATTCTGAATTCGTGAAACCGCTCAACAAGGCTGTTGGCTCAAGTGTCAAGGAAAT	286
Db	107	CTTGCAACATTGGGGATGCTGAGGCGAGTGAAGCTCACTTCAGCTCCATCAACGACACC	166
OY	287	TTGGTAGTGGCCATTGCACTTGGTTAAACACAGCTGGTTACTGTGAAAACTTCCCATGTG	346
Db	167	ACGGCAAGATGCTGACTGTTGATTTAACACCGCTGGATTACTGTGAAAACTTCCCTGCGG	226
OY	347	AAGATTATCCCAAGCAAGAACGCTGAGAGATGCTGAAGGTTAACTTGTGGGTTCTTTGT	406
Db	227	AAACGTACCCGGCTACTAAACGCTGAAGCATATGAAGGTGAACGGTTTGGGCTCATTTCT	286
OY	407	ATGTTTCCAAACCTTGTGACCAATGATCAAAAGATATCAAGGGGCTTCGTG	466
Db	287	ACGTTTGGAAATCGTTGCTGAACCAATTATCCAGAA---CAATCGAGAGCTCTATCA	343
OY	467	TTTTGATTGGTCTATATGCTGGGCAATGTGCACAGATCCTCAAAACCAAGTTGTCTACA	526
Db	344	TCTTGATTTGGCTCANTGTCTGAAACAATTGTCAACGACCAACCAATGATATGTACA	403
OY	527	ACATGTCCAAAGCTGTGTTATTCATTGAGCTTGAAGACTTTGGCTTGTGAATGGGCTTAAGT	586
Db	404	ACATGTCCAAAGCTGTGAGTGAATCCACTTGGTCAAGTCGTGGGCTCGAATGGGCCAAGT	463
OY	587	ACAAATCAAGATTAATTCCTTAAACCCAGGTTACATCAAGGTCCTTTGAACCAAGAAATG	646
Db	464	ACAAATCAAGATCAACACTTATCAACAGGCTATATTTTGACTCCTTTTACCGAAGAACG	523
OY	647	TTATCAATGTGTACAGAAATGTGTACAAACAGATGATCTCTGGTATCCCAACAAGAA	706
Db	524	TGATTTTCTGGCCACACAGAGATGAAAGGAAGCTGGGAATCCAAAGATCCCATGAAGAA	583
OY	707	TGTCCGAACCAAGGAATAATTGGTGTGTTTGTACTTGTCTTGAATCTGTCTG	766
Db	584	TGGCCGAATCCCAAGGAATTCGTGGGGGTCCATCTTATACTGTGCAAGCGAGACTGTTCTT	643
OY	767	CATACACTATCTGTGCCACTTACTGTTGATATGATGTTG	805
Db	644	CCACACTTACGGGCCCAATTTGGTTGTGTGACCGAAGAT	682

RESULT 2
US-09-184-965-6
: Sequence 6, Application US/09184965
: Patent No. 6582944
: GENERAL INFORMATION:
: APPLICANT: Halborn, Johan
: APPLICANT: Penttila, Merja
: APPLICANT: Ojamo, Heikki
: APPLICANT: Keranen, Sirka
: APPLICANT: Hahn-Hagerdal, Barbel
: APPLICANT: Waldfriðsson, Mats
: APPLICANT: Aitakainen, Ulla
: TITLE OF INVENTION: Xylose Utilization by Recombinant Yeasts
: NUMBER OF SEQUENCES: 9
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Birch, Stewart, Kolaasch & Birch
: STREET: 301 N. Washington St.
: CITY: Falls Church
: STATE: Virginia
: COUNTRY: USA

```

1 ZIP: 22046-3487
2
3 COMPUTER READABLE FORM:
4
5 MEDIUM TYPE: Floppy disk
6
7 COMPUTER: IBM PC compatible
8
9 OPERATING SYSTEM: PC-DOS/MS-DOS
10
11 SOFTWARE: Patentin Release #1.0, Version #1.25
12
13 CURRENT APPLICATION DATA:
14
15 APPLICATION NUMBER: US/09/184,965
16
17 FILING DATE:
18
19 CLASSIFICATION:
20
21 PRIOR APPLICATION DATA:
22
23 APPLICATION NUMBER: US/08/336,198
24
25 FILING DATE: 03-NOV-1994
26
27 ATTORNEY/AGENT INFORMATION:
28
29 NAME: Murphy Jr., Gerald M.
30
31 REGISTRATION NUMBER: 28,977
32
33 TELECOMMUNICATION INFORMATION:
34
35 TELEPHONE: 703-205-8000
36
37 TELEFAX: 703-205-8050
38
39 TELEX: 248345
40
41 INFORMATION FOR SEQ ID NO: 6:
42
43 SEQUENCE CHARACTERISTICS:
44
45 LENGTH: 696 base pairs
46
47 TYPE: nucleic acid
48
49 STRANDEDNESS: double
50
51 TOPOLOGY: linear
52
53 MOLECULE TYPE: cdna
54
55 HYPOTHEICAL: NO
56
57 ORIGINAL SOURCE:
58
59 ORGANISM: Pichia stipitis
60
61 STRAIN: CBS-6054
62
63 FEATURE:
64
65 NAME/KEY: CDS
66
67 LOCATION: 1..693
68
69 OTHER INFORMATION: /standard_name="xyliitol"
70
71 OTHER INFORMATION: dehydrogenase"
72
73 PUBLICATION INFORMATION:
74
75 DOCUMENT NUMBER: FI 901771
76
77 FILING DATE: 06-APR-1990
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968

```

Query Match	29.2%	Score 238.2	DB 4	Length 696
Best Local Similarity	64.4%	Pred. No. 9.4e-65		
Matches 373	Conservative 0	Mismatches 203	Indels 3	Gaps 1
QY	227	CCTGGAATATTTCGATTCGTATACCGTTCACAAGAGTGTCTCAAGTTGCTAAGATT	286	
Db	107	CCTGGAACATTGGGGATCTGAGGACGTAAAGCGTACTTTCAGTCCATCAACGAACACC	166	
QY	287	TTGGTAAGTTGCCATTCGACTTTGGTTAACACAGCTGTCTACTGTGAAACCTTCCATGTG	346	
Db	167	AACGGGAATGCTGCACTGTGTGATTAACACCGCTGGAACTGTGAAACCTTCCGCG	226	
QY	347	AAGATTACCCAGCCAAAGACGCTGAGAGATGTGTAAGTTAACTGTGTGGTTCTTGT	406	
Db	227	AAACCTACCCGGCTACTTAACCTTAAAGCATCATGAAGTAAAGGTTTGGGCTCATCT	286	
QY	407	ATGTTTCTCAAGCCTTGTCTTAAGCCATTGATTCAAAGAAAGGTATCAAGGCTGCTTCTGTG	466	
Db	287	ACGTTTCCCAATCGTTCCTAGACCATTGATTCAGAA---CACTGTGAGAGCCTTTATCA	343	
QY	487	TTTGTATTGGTTCATGTCTGTGTCATGTTCACAGATCTCAAAACCAAGTTGTCAC	526	
Db	344	TTTGTATTGGCTCAATGTCTGTGAACAATTTGTTCACGACCCACMACCCCAATGTATGTCA	403	
QY	527	ACATGTCCAAAGCTGTGTTATTCATTGTGGCTAAGACTTGTGGCTGTGAATGGCTAAGT	586	
Db	404	ACATGTCCAAAGCTGTGAATGATTCACCTGTGTCAATCGTTGGCCCTGGCAATGGGCCAAGT	463	
QY	587	ACAAATACAGATTATTTCTTAAACCCAGATACATCTACGGTCTTTGACCAAGATG	646	
Db	464	ACAAATACAGATCAACCTTATCACCAAGGCTATATTTTGAATCTCTTTAACCAAGAACG	523	

QY 647 TTATCATGTGTAAGAAATTTGACACAGATGATCTGTGATATCCCAACAAGAA 706
DB 524 TATATTTCTGGCCACACAGAGATGAAGAGCCCTGGAAATCCAGATCCCATGAAAGAA 583
QY 707 TGTCCGAACCAAGAAATATCTGTGCTGTGTTTGTATCTGTTCTGATGCTGCT 766
DB 584 TGGCCGAACCAAGAAATCTGTGGGTCCATCTTATCTGGCAAGCGAGATGCTTCTT 643
QY 767 CATACACTACTGTGTCAGCTTACTGTGTTGATGTTGTT 805
DB 644 CCTACACTAGGGGCCCAATTTGTTGTGACGAGAT 682

RESULT 3
US-09-248-796A-29

/ Sequence 29, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstein et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ PRIOR FILING DATE: 1998-02-12
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ NUMBER OF SEQ ID NOS: 28208
/ LENGTH: 846
/ TYPE: DNA
/ ORGANISM: Candida albicans
US-09-248-796A-29

Query Match 26.8%; Score 218.6; DB 4; Length 846;
Best Local Similarity 62.0%; Pred. No. 1.6e-58;
Matches 385; Conservative 0; Mismatches 224; Indels 12; Gaps 2;

QY 14 TTCCAATTTTATGATTCGATGCGCACTTAACCATGTGACAGGTGCTGTGGTTAG 73
DB 218 TCCCAAGCTTCGTTGATGAGAAACATGATTTAACCGTGGCTGTGGTTAG 277
QY 74 CTGAAGCTTTATCAAGGGTTGTTGCTGCTACGTTCTGACATTCCTTGTCTGATTCG 133
DB 278 CTGCGGTGTATCAAGAGCTTTATTAAGCAAGGTCGATGTTGCTATTAAGCATTA 337
QY 134 ACCAAGAAAGATCTCTGCAAGCCGATACCAAAATAGCTACTGAAGATTGA 193
DB 338 ACTTGAAAGAACACACAAAGCTGTAGAGCTTTACAAATGGGCGAAGAGCAAAATGA 397
QY 194 AGTGAAGAGATTCCTCA-----AGATGGTTTCAATGCTGTGATTTCTGAT 244
DB 398 AAGGTAATTCGAATCAACCAATCGTCAAGTGAAGTCTTGTGATTAATATGCGATG 457
QY 245 CTGATACGTTCAAGGTGTTTGTCTCAAGTTGCTAAGATTTTGTGAAGTCCATTGC 304
DB 458 CTGAAGCTGTGCACTTGAATTCAGAAAGCATCAAGAAACCAAGGCAAAATCTCAAG 517
QY 305 ACTTGTTAAACAGCTGTTACTGTGAATACTTCCATGTAAGATTACCAAGCAAGA 364
DB 518 TCTTGTTCAACATGCGCGTTACGCTGAACCTTCCAGCTGAAAGTAACCAAGCAAGA 577
QY 365 ACGCTGAGAAGATGTAAGTTTACTTGGGTTCTTTGTATTTCTCAAGCTTTG 424
DB 578 ACGCTGAAGATTTATGAAGTTTACGAGTTGGGTTCTTCTCAAGCTTTTGG 637
QY 425 CTAAACCATGTATCAAGAGATTAACAGGTGCTTCTGTGTTTGTATGTTGTTATGT 484
DB 638 CTAGACCATTTATCCAAA-----CAACATGACCGGATCATTTTGAATGGGTAAAT 694
QY 485 CTGGTGCATTTGTCAAGATCTCTCAAAACCAAGTTGTCTACACATGTCAGAGCTGTG 544

DB 695 CCGTACCATGCTACACACCAACCAACCAATGATGTAACATGTCACAAAGCCGGTG 754
QY 545 TTATCATTTGGCTAAGCTTTGGCTTGTGATGAGGCTTAAGTAACATCAGATTAA 604
DB 755 TCATTCATTTAGCCAGATCATTTGGCTGTGAATGGCTTAATACATATCAAGTCA 814
QY 605 CTTTAAACCCAGTTTACATCT 625
DB 815 CATTTGTCGCGCGGACATCT 835

RESULT 4
US-09-248-796A-2490

/ Sequence 2490, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstein et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ PRIOR FILING DATE: 1998-02-12
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ NUMBER OF SEQ ID NOS: 28208
/ LENGTH: 927
/ TYPE: DNA
/ ORGANISM: Candida albicans
US-09-248-796A-2490

Query Match 10.6%; Score 86.6; DB 4; Length 927;
Best Local Similarity 52.0%; Pred. No. 8.1e-17;
Matches 328; Conservative 0; Mismatches 279; Indels 24; Gaps 5;

QY 182 CTGAAGATTTGAAGTTGAAGAAAGTTCCAAAGATGGGTTCAATAGCTGTGATATTTCTG 241
DB 308 CTGATATTTTACCAAGACTTATGGGTTAATCCAAAGCATCAAAATGATATTTACTG 367
QY 242 ATTCTGATACGTTCAAGGTGTTTGTCTCAAGTTGCTAAGATTTTGTGAAGTTGCAT 301
DB 368 ATTCCAAAGATTTGAAGAAAGTTGTACAAACAAATGAATGATGATTTGCGTACATGATA 427
QY 302 TGCACTGTGTTAAACAGCTGTGATCTG--TGAAACTTCCATGTAAATTAACCAAG 358
DB 428 TCTTGTGTCCAAGTCTGTGTTGCTGTGACCGAAGGCTCGAATGATGTCAAGGAG 487
QY 359 CCAAGAACGCTGAGAAGATGTAAGTTTACTTGTGGTTCTTTGTATGTTTCTCAAG 418
DB 488 TCGAACAATGAAACAAAGTTTGTATGTATTTAAACAGTGTATTTATGTCATG 547
QY 419 CTTTGTCAAGCATTTGATCAAGAAAGATTAACAGGTGCTTCTGTGTTTGTATGTTG 478
DB 548 TTGTTGTCAATTTTCAAGAAAGGTA-----AAGGTCATTCATTTTCACTGCA 601
QY 479 CTATGCTGTGTCATTTGTCAAGATCTCTCAAAACCAAGTTGTCTCAACATGTCAGG 538
DB 602 GTATGTGGCTTCAATTTGATATGTCCCAATTAACAGCAGCTTCAACCTGTAAAG 661
QY 539 CTGAGTTATTCATTTGGCTAAGACTTGTGCTTGTGATGAGGCTAAGTAACATCAAG 598
DB 662 CTGGGCTCAACATTTGTCCAAATCATTTGATGTGTAATGGGCAACATTTGCT--AGAG 718
QY 599 TTAATTTTAAACCAAGTTTACATCAAGTCTCTTGAACCAAGATGTTTACATAGTA 658
DB 719 TCAATTTCTGTTTCTCAAGTTTACATGCTACTCATTTAGATTTGCTG----- 769
QY 659 ACGAAGATTTGTAACAGATGATCTCTGTATCCCAACAAAGATGTCGAACCA 718
DB 770 ATCCGATGTCAGAGTAATGTTGTTGCACTTACACACTTGTGTAGAGAACCAACCA 829

RESULT 6
US-09-248-796A-2654
; Sequence 2654, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:

```

RESULT 7
US-09-248-796A-2488
/ Sequence 2488, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keltch weinstock et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANCA
/ TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248, 796A
/ CURRENT FILING DATE: 1999-02-12
/ PRIOR APPLICATION NUMBER: US 60/074, 725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096, 409
/ PRIOR FILING DATE: 1998-08-13

```

NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 2488
LENGTH: 336
TYPE: DNA
ORGANISM: Candida albicans
US-09-248-796A-2488

Query Match
Best Local Similarity 52.9%; Score 55; DB 4; Length 336;
Matches 171; Conservative 0; Mismatches 140; Indels 12; Gaps 2;

QY 491 CCATTGTCACGATCTCTCAAAAACCAAGTGTCTACAAACATGTCACAGGCTGGGTATCC 550
DB 20 CCGGTGTGAATACCTCAATTAACAAGCTCTTATATCTGTCTAAGCTGATGACTC 79
QY 551 ATTGGCTTAAGCTTTGGCTTGTGAATGGCTTAAGTACCAATCAGATTAATCTTTAA 610
DB 80 ATTATGCCAATCATTTAGAGTGAATGGCTAGTTTGTGCTAGAGTAATTCATTT 139
QY 611 ACCGAGTTACATCAAGCTCCTTTGACCAAGATTTATCAATGTATACGAAGATTGT 670
DB 140 CTCACGGGTATATA-----TTGACTGATATTTGCTGATTCAGAAATGA 190
QY 671 ACAACAGATGATCTCTGTATCCCAACAAGATGTCGACCAAGAAATACATTG 730
DB 191 AAAAATATGTGCAATTAACCTTTGGAGAGAGATTAACCAAGAAATTAAGTGG 250
QY 731 GTGCTTTTGTACTTGTCTTGAATCTGTCTTACATACATACGCTGCGCACTTAC 790
DB 251 GGGCATTTTATTACTTGGCTC---AAATGATCAACTATTAATCTGTTCAATATTG 307
QY 791 TGGTGAATGCTGTTTCACTTCT 813
DB 308 CTGTTGATGGGGTATATCATGT 330

RESULT 8

US-09-248-796A-6987/C
Sequence 6987, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
PRIOR FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 6987
LENGTH: 372
TYPE: DNA
ORGANISM: Candida albicans
US-09-248-796A-6987

Query Match
Best Local Similarity 6.5%; Score 53; DB 4; Length 372;
Matches 83; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 14 TTCCACTTTAGATTCGATGCGCACTTAACCATTTGCACAGTGCCTGTGGTGTAG 73
DB 205 TCCCAAGCTTCGGTTGATGGAAGAACTAGTATTTAACCGGTGGCTCTGTGGTTGG 146
QY 74 CTGAAGCTTTATCAAGGCTTTGGCTTACGCTTACAGTTCCTTGTATATCG 133
DB 145 CTGCGGTGTATCAAGACTTTATTAAGCAAGGTGCGAGTGTGCATTAGTCATATGA 86
QY 134 ACCAAGAAAGAC 146
DB 85 ACTTGAAAGAAC 73

RESULT 9
US-09-134-001C-725
Sequence 725, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 725
LENGTH: 774
TYPE: DNA
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-725

Query Match
Best Local Similarity 47.8%; Score 53; DB 3; Length 774;
Matches 223; Conservative 0; Mismatches 235; Indels 9; Gaps 2;

QY 200 AAGAATCCAAAGATGGCTCATATGCTGTGATTTCTGATTCGATACCGTCA 259
DB 182 AATCAGGTGTTAAACCAATCATTAATAATTAAGATTAATCACTTAATGAAGTAATC 241
QY 260 AGGTGTTGCTCAAGTGTCTAAGATTTTGGTAAAGTCCATTCGACTTGTAAACAG 319
DB 242 AATTTGTTGATTAATGTAAGGAATATGAAATATGATATTC--TTGTCATTAACG 298
QY 320 CTGTTACTGTGAACACTCCCATGTGAAGATTACCCAGCAAGCCGTAAGAAATG 379
DB 299 CTAGATCTCAATTCAGATGATCTGAATAATTTTGTATGAAGATGTTAAAGAA 358
QY 380 TGAAGTTAACTTGTGGGTTCTTTGATTTTCTCAAGCCTTGTCTAAGCATGATCA 439
DB 359 TCAATTAAGCTTAATGAGCAATTCCTGTAGGCAAAAGCTGAGCAATTAATGATG 418
QY 440 AAGAAGTATCAAGGCTCTGTGTTGTTGATTTGATGTTGCTATGCTGTCGCAATTGCA 499
DB 419 AGAAGGCA-----GTGTTGATGATTAATGCTCTTCACTTATGAGATTGATGCTA 472
QY 500 ACGATCTCAAAACCAAGTGTCTAACAATGTCCAAGGCTGTGTTATTCATTGGCTA 559
DB 473 ATAAAGCAAGACAGAGCTCTTATGAAGCTTAAGCCGGTGTGACATGCTGACTA 532
QY 560 AGACTTGGCTTGTGAATGGCTTAAGTACCAATCAGATTAATCTTTAAACCAAGTT 619
DB 533 AAAGCTTAGCTAAGAAATGCTGATATGAATTAAGTAATGAATGCAATGCCCCGTGTT 592
QY 620 ACATCTACGTCCTTTGACCAAGATGTTATCATGTAAGCAAGAA 666
DB 593 ACATGAAACATCAAGAAAGCAAGTAATTAATGATTAATCTGAA 639

RESULT 10

US-09-248-796A-2730
Sequence 2730, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
PRIOR FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725

QY 14 TTCCACTTTAGATTCGATGCGCACTTAACCATTTGCACAGTGCCTGTGGTGTAG 73
DB 205 TCCCAAGCTTCGGTTGATGGAAGAACTAGTATTTAACCGGTGGCTCTGTGGTTGG 146
QY 74 CTGAAGCTTTATCAAGGCTTTGGCTTACGCTTACAGTTCCTTGTATATCG 133
DB 145 CTGCGGTGTATCAAGACTTTATTAAGCAAGGTGCGAGTGTGCATTAGTCATATGA 86
QY 134 ACCAAGAAAGAC 146
DB 85 ACTTGAAAGAAC 73

Query Match	6.3%;	Score 51.2;	DB 3;	Length 852;
Best Local Similarity	58.6%;	Pred. No. 1.2e-05;		
Matches	89;	Conservative	0;	Matches 63; Indels 0; Gaps 0
Qy	437	TCAAAGAAGTATCAAGGGTGGCTTCGTGGTTTGTATTGATGGTCTCATATGCTGGTGCATTTG	496	
Db	485	TCAGAGAAGAGGCGCAAGAGGGCGCCCTGTGTTCACGGCTTCATGTCTGGCCACATTTG	544	
Qy	497	TCACAGATCTCAAAAACCAAGTTGTCTACACATGTCCAAAGCGTGGTATTATCAATTTGG	556	
Db	545	TGAACCTGCCCAAGTTCCAGGCCACGTACAAAGCGGCCCAAGGCTGGCGTGGCCCACTTGG	604	
Qy	557	CTAAGACTTTGGCTGTGTGAATGGCTTAAGTAC	588	

Db 605 CGAAGTCCGTGCGCCGTCAGATTCCGCCGCTTC 636

RESULT 13
US-09-734-237B-69

; Sequence 69, Application US/09734237B
; Patent No. 6818752
; GENERAL INFORMATION:
; APPLICANT: Rozzell, J. David
; APPLICANT: Bul, Peter
; APPLICANT: Hua, Ling
; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
; FILE REFERENCE: B583:40608
; CURRENT APPLICATION NUMBER: US/09/734,237B
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/494,921
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patent version 3.1
; SEQ ID NO 69
; LENGTH: 852
; TYPE: DNA
; ORGANISM: Candida magnoliae
US-09-734-237B-69

Query Match 6.3%; Score 51.2; DB 4; Length 852;
Best Local Similarity 58.6%; Pred. No. 1.2e-05;
Matches 89; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 437 TCAAGAAGATATCAAGGCTGCTTGTGTTTATGTTCTATGCTGTGTCATTTG 496
Db 485 TCGAAGAGAGGCGCAAGAGGCGCCCTTGTGTTTCAAGCCCTTCATGCTGCCACATTTG 544
Qy 497 TCAAGATCTCTCAAAACCAAGTTGTCTAACAACATGTCACAGGCTGTGTTATTCATTTGG 556
Db 545 TGAAGTCCGCCAGTTCAGGCGCACGTACACGCGCCCAAGGCTGCGCTGCCCATTTGG 604
Qy 557 CTAAAGCTTTGGCTTGTGATGAGGCTTAAGTAC 588
Db 605 CGAAGTCCGTGCGCCGTCAGATTCCGCCGCTTC 636

RESULT 14
US-09-489-039A-5819

; Sequence 5819, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709,2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 5819
; LENGTH: 951
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-5819

Query Match 6.3%; Score 51.2; DB 4; Length 951;
Best Local Similarity 54.9%; Pred. No. 1.2e-05;
Matches 101; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

Qy 454 GGTGCTCTGTTGTTTATGTTTATGTTGTCGTCATGTCACAGATCTCAAAAC 513
Db 595 GCGGATGATTAATCAATATCGCTCGATGTCGGGATTAATTTAAACGCGGCTCGAC 654
Qy 514 CAAATTGCTACACATATCCAGGCTGTGTATCCATTTGGCTTAAGATTTGGCTTGT 573

Db 655 CAGGCGATTAACAATGCTCCAAAGCGCGGTATCATCTGTCCAAAAAGCTGCGATG 714
Qy 574 GAATGGGCTAAGTACACATCAAGTAATTTCTTAAACCAAGTATCAATCAAGTCTCT 633
Db 715 GAGTGGGTGGCAAGAGGATCCGCTCAATTCATCAATCCGGGATATACCGCACGCGG 774
Qy 634 TTGA 637
Db 775 ATGA 778

RESULT 15
US-09-367-012-8

; Sequence 8, Application US/09367012
; Patent No. 6218156
; GENERAL INFORMATION:
; APPLICANT: Yasohara, Yoshihiko
; APPLICANT: Kizaki, No. 6218156iyuki
; APPLICANT: Hasegawa, Junzo
; APPLICANT: Wada, Masaru
; APPLICANT: Shimizu, Sakayu
; APPLICANT: Kataoka, Michihiko
; APPLICANT: Yamamoto, Kazuhiko
; APPLICANT: Kawabata, Hiroshi
; TITLE OF INVENTION: No. 6218156el carbonyl reductase, gene coding same, and method
; FILE REFERENCE: S72030PUS
; CURRENT APPLICATION NUMBER: US/09/367,012
; CURRENT FILING DATE: 1999-11-24
; EARLIER APPLICATION NUMBER: PCT/JP97/03051
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 8
; LENGTH: 1303
; TYPE: DNA
; ORGANISM: Candida magnoliae
US-09-367-012-8

Query Match 6.3%; Score 51.2; DB 3; Length 1303;
Best Local Similarity 58.6%; Pred. No. 1.5e-05;
Matches 89; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 437 TCAAGAAGATATCAAGGCTGCTTGTGTTTATGTTTATGTTCTATGCTGTGTCATTTG 496
Db 643 TCGAAGAGAGGCGCAAGAGGCGCCCTTGTGTTTCAAGCCCTTCATGCTGCCACATTTG 702
Qy 497 TCAAGATCTCTCAAAACCAAGTTGTCTAACAACATGTCACAGGCTGTGTTATTCATTTGG 556
Db 703 TGAAGTCCGCCAGTTCAGGCGCACGTACACGCGCCCAAGGCTGCGCTGCCCATTTGG 762
Qy 557 CTAAAGCTTTGGCTTGTGATGAGGCTTAAGTAC 588
Db 763 CGAAGTCCGTGCGCCGTCAGATTCCGCCGCTTC 794

Search completed: July 12, 2005, 15:30:36
Job time : 191 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: July 12, 2005, 13:24:26 ; Search time 638 Seconds

(without alignments)
8029.958 Million cell updates/sec

Title: US-10-720-018-1

Perfect score: 816
Sequence: 1 atgcactgactacattccacac.....atggtgttcttcttcttg 816

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 6330945 seqs, 3139162390 residues

Total number of hits satisfying chosen parameters: 12661690

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*\n2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*\n3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*\n4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*\n5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*\n6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*\n7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*\n8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*\n9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*\n10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*\n11: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*\n12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*\n13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*\n14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*\n15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*\n16: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*\n17: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*\n18: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*\n19: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*\n20: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*\n21: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*\n22: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*\n23: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*\n24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*\n25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*\n26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	816	100.0	816	19	US-10-720-018-1
2	99.4	12.2	3203	20	US-10-653-047-41
3	88.6	10.9	687	20	US-10-653-047-4932
4	81.8	10.0	637	17	US-10-369-493-36723
5	80.6	9.9	735	17	US-10-369-493-26991
6	74.6	9.1	852	10	US-09-734-2378-71
7	74.2	9.1	1146	20	US-10-653-047-4825

8	72.4	8.9	775	20	US-10-425-115-172450	Sequence 172450, A
9	67.6	8.3	783	17 <th>US-10-369-493-25810</th> <th>Sequence 25810, A</th>	US-10-369-493-25810	Sequence 25810, A
10	63.6	7.8	774	17 <th>US-10-369-493-26610</th> <th>Sequence 26610, A</th>	US-10-369-493-26610	Sequence 26610, A
11	60.6	7.4	636	20 <th>US-10-653-047-3866</th> <th>Sequence 3866, Ap</th>	US-10-653-047-3866	Sequence 3866, Ap
12	59.8	7.3	822	15 <th>US-10-314-394-1</th> <th>Sequence 1, Appli</th>	US-10-314-394-1	Sequence 1, Appli
13	56.4	6.9	825	15 <th>US-10-314-394-3</th> <th>Sequence 3, Appli</th>	US-10-314-394-3	Sequence 3, Appli
14	56.2	6.9	789	9 <th>US-09-938-842A-948</th> <th>Sequence 948, App</th>	US-09-938-842A-948	Sequence 948, App
15	55.2	6.8	732	17 <th>US-10-369-493-38959</th> <th>Sequence 38959, A</th>	US-10-369-493-38959	Sequence 38959, A
16	55.2	6.8	732	17 <th>US-10-369-493-38960</th> <th>Sequence 38960, A</th>	US-10-369-493-38960	Sequence 38960, A
17	55.2	6.8	837	20 <th>US-10-425-115-100705</th> <th>Sequence 100705, A</th>	US-10-425-115-100705	Sequence 100705, A
18	55.2	6.8	949	9 <th>US-09-770-445-339</th> <th>Sequence 339, App</th>	US-09-770-445-339	Sequence 339, App
19	55.2	6.7 <td>783</td> <td>9<th>US-09-938-842A-1025</th><th>Sequence 1025, App</th></td>	783	9 <th>US-09-938-842A-1025</th> <th>Sequence 1025, App</th>	US-09-938-842A-1025	Sequence 1025, App
20	54.4	6.7 <td>783</td> <td>11<th>US-09-938-842A-1025</th><th>Sequence 1025, Ap</th></td>	783	11 <th>US-09-938-842A-1025</th> <th>Sequence 1025, Ap</th>	US-09-938-842A-1025	Sequence 1025, Ap
21	53.8	6.6 <td>768</td> <td>17<th>US-10-369-493-26631</th><th>Sequence 26631, A</th></td>	768	17 <th>US-10-369-493-26631</th> <th>Sequence 26631, A</th>	US-10-369-493-26631	Sequence 26631, A
22	53.4	6.5 <td>852</td> <td>19<th>US-10-639-159-53</th><th>Sequence 53, Appl</th></td>	852	19 <th>US-10-639-159-53</th> <th>Sequence 53, Appl</th>	US-10-639-159-53	Sequence 53, Appl
23	53.2	6.5 <td>852</td> <td>19<th>US-10-639-159-53</th><th>Sequence 53, Appl</th></td>	852	19 <th>US-10-639-159-53</th> <th>Sequence 53, Appl</th>	US-10-639-159-53	Sequence 53, Appl
24	53.2	6.5 <td>852</td> <td>20<th>US-10-782-258-53</th><th>Sequence 53, Appl</th></td>	852	20 <th>US-10-782-258-53</th> <th>Sequence 53, Appl</th>	US-10-782-258-53	Sequence 53, Appl
25	53.2	6.5 <td>852</td> <td>20<th>US-10-782-258-53</th><th>Sequence 53, Appl</th></td>	852	20 <th>US-10-782-258-53</th> <th>Sequence 53, Appl</th>	US-10-782-258-53	Sequence 53, Appl
26	53.2	6.5 <td>855</td> <td>19<th>US-10-639-159-45</th><th>Sequence 45, Appl</th></td>	855	19 <th>US-10-639-159-45</th> <th>Sequence 45, Appl</th>	US-10-639-159-45	Sequence 45, Appl
27	53.2	6.5 <td>855</td> <td>19<th>US-10-639-159-45</th><th>Sequence 45, Appl</th></td>	855	19 <th>US-10-639-159-45</th> <th>Sequence 45, Appl</th>	US-10-639-159-45	Sequence 45, Appl
28	53.2	6.5 <td>855</td> <td>20<th>US-10-782-258-39</th><th>Sequence 39, Appl</th></td>	855	20 <th>US-10-782-258-39</th> <th>Sequence 39, Appl</th>	US-10-782-258-39	Sequence 39, Appl
29	53.2	6.5 <td>855</td> <td>20<th>US-10-782-258-39</th><th>Sequence 39, Appl</th></td>	855	20 <th>US-10-782-258-39</th> <th>Sequence 39, Appl</th>	US-10-782-258-39	Sequence 39, Appl
30	53.2	6.5 <td>855</td> <td>20<th>US-10-782-258-47</th><th>Sequence 47, Appl</th></td>	855	20 <th>US-10-782-258-47</th> <th>Sequence 47, Appl</th>	US-10-782-258-47	Sequence 47, Appl
31	52.8	6.5 <td>819</td> <td>15<th>US-10-259-165-699</th><th>Sequence 699, App</th></td>	819	15 <th>US-10-259-165-699</th> <th>Sequence 699, App</th>	US-10-259-165-699	Sequence 699, App
32	52.4	6.4 <td>656</td> <td>9<th>US-09-770-149-497</th><th>Sequence 497, App</th></td>	656	9 <th>US-09-770-149-497</th> <th>Sequence 497, App</th>	US-09-770-149-497	Sequence 497, App
33	52.2	6.4 <td>852</td> <td>19<th>US-10-639-159-51</th><th>Sequence 51, Appl</th></td>	852	19 <th>US-10-639-159-51</th> <th>Sequence 51, Appl</th>	US-10-639-159-51	Sequence 51, Appl
34	52.2	6.4 <td>852</td> <td>20<th>US-10-782-258-51</th><th>Sequence 51, Appl</th></td>	852	20 <th>US-10-782-258-51</th> <th>Sequence 51, Appl</th>	US-10-782-258-51	Sequence 51, Appl
35	51.6	6.3 <td>852</td> <td>19<th>US-10-639-159-55</th><th>Sequence 55, Appl</th></td>	852	19 <th>US-10-639-159-55</th> <th>Sequence 55, Appl</th>	US-10-639-159-55	Sequence 55, Appl
36	51.6	6.3 <td>852</td> <td>19<th>US-10-639-159-55</th><th>Sequence 55, Appl</th></td>	852	19 <th>US-10-639-159-55</th> <th>Sequence 55, Appl</th>	US-10-639-159-55	Sequence 55, Appl
37	51.6	6.3 <td>852</td> <td>20<th>US-10-782-258-55</th><th>Sequence 55, Appl</th></td>	852	20 <th>US-10-782-258-55</th> <th>Sequence 55, Appl</th>	US-10-782-258-55	Sequence 55, Appl
38	51.6	6.3 <td>852</td> <td>20<th>US-10-782-258-55</th><th>Sequence 55, Appl</th></td>	852	20 <th>US-10-782-258-55</th> <th>Sequence 55, Appl</th>	US-10-782-258-55	Sequence 55, Appl
39	51.6	6.3 <td>852</td> <td>20<th>US-10-782-258-113</th><th>Sequence 113, Appl</th></td>	852	20 <th>US-10-782-258-113</th> <th>Sequence 113, Appl</th>	US-10-782-258-113	Sequence 113, Appl
40	51.6	6.3 <td>852</td> <td>20<th>US-10-782-258-113</th><th>Sequence 113, Appl</th></td>	852	20 <th>US-10-782-258-113</th> <th>Sequence 113, Appl</th>	US-10-782-258-113	Sequence 113, Appl
41	51.6	6.3 <td>855</td> <td>19<th>US-10-639-159-43</th><th>Sequence 43, Appl</th></td>	855	19 <th>US-10-639-159-43</th> <th>Sequence 43, Appl</th>	US-10-639-159-43	Sequence 43, Appl
42	51.6	6.3 <td>855</td> <td>20<th>US-10-782-258-43</th><th>Sequence 43, Appl</th></td>	855	20 <th>US-10-782-258-43</th> <th>Sequence 43, Appl</th>	US-10-782-258-43	Sequence 43, Appl
43	51.4	6.3 <td>849</td> <td>16<th>US-10-032-585-6746</th><th>Sequence 6746, App</th></td>	849	16 <th>US-10-032-585-6746</th> <th>Sequence 6746, App</th>	US-10-032-585-6746	Sequence 6746, App
44	51.2	6.3 <td>711</td> <td>15<th>US-10-259-165-770</th><th>Sequence 770, App</th></td>	711	15 <th>US-10-259-165-770</th> <th>Sequence 770, App</th>	US-10-259-165-770	Sequence 770, App
45	51.2	6.3 <td>810</td> <td>9<th>US-09-938-842A-1690</th><th>Sequence 1690, App</th></td>	810	9 <th>US-09-938-842A-1690</th> <th>Sequence 1690, App</th>	US-09-938-842A-1690	Sequence 1690, App

ALIGNMENTS

RESULT 1
US-10-720-018-1
; Sequence 1, Application US/10720018
; Publication No. US20040132074A1
; GENERAL INFORMATION:
; APPLICANT: Verho, Rltva
; APPLICANT: Richard, Peter
; APPLICANT: Penttila, Merja
; TITLE OF INVENTION: New Enzyme for an in vivo and in vitro utilisation of
; FILE REFERENCE: 2530-120
; CURRENT APPLICATION NUMBER: US/10/720, 018
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 816
; TYPE: DNA
; ORGANISM: Ambrosiozyma monospora
US-10-720-018-1

Query Match 100.0%; Score 816; DB 19; Length 816;
Best Local Similarity 100.0%; Pred. No. 2.4e-230;
Matches 816; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 ATGACTGACTACATTCCAGCTTTAGTGGAGCCATTACCTTGCAGGTCGCC 60
Db 1 ATGACTGACTACATTCCAGCTTTAGTGGAGCCATTACCTTGCAGGTCGCC 60
Oy 61 TGTGTTGTTAGCTGAGCTTATCATCAGGCTTTGTGCTGCTGCTGCTGCT 120

```
Db 61 TGTGTGTTAGTGAAGCTTATCAAGGTTTGTGGCTAGGTTGTCATTTGCT 120
Qy 121 TTGCTGATATGACCAAGAAAGAGCTGCTCCCAACAGCCGAATACCAATAGCT 180
Db 121 TTGCTGATATGACCAAGAAAGAGCTGCTCCCAACAGCCGAATACCAATAGCT 180
Qy 181 ACTGAAGATGAAAGTTGAAGAGTTCCCAAGATGGGTTCAATGCGCTGATATTTCT 240
Db 181 ACTGAAGATGAAAGTTGAAGAGTTCCCAAGATGGGTTCAATGCGCTGATATTTCT 240
Qy 241 GATTCTGATACCTGTCACAGAGTGTGCTCAAGTTCAGTGAAGATTTTGTAGTGGCA 300
Db 241 GATTCTGATACCTGTCACAGAGTGTGCTCAAGTTCAGTGAAGATTTTGTAGTGGCA 300
Qy 301 TTGCACTGTTTAAACAAGCTGCTGTTACGTGAAACCTCCATGGAAGATTAACCCAGC 360
Db 301 TTGCACTGTTTAAACAAGCTGCTGTTACGTGAAACCTCCATGGAAGATTAACCCAGC 360
Qy 361 AAGAAGCTGAGAGATGATGAGGTTAACTTGTGGTCTTGTGATGTTTCTCAAGCC 420
Db 361 AAGAAGCTGAGAGATGATGAGGTTAACTTGTGGTCTTGTGATGTTTCTCAAGCC 420
Qy 421 TTGCTAAGCCATTGATCAAGAGATCAAGAGGTCCTTCTGTTTGAATGTTCT 480
Db 421 TTGCTAAGCCATTGATCAAGAGATCAAGAGGTCCTTCTGTTTGAATGTTCT 480
Qy 481 ATGCTGCTGCTATGTCACAGTCTCTCAAAACCAAGTTCACAAATGTCAGAGCT 540
Db 481 ATGCTGCTGCTATGTCACAGTCTCTCAAAACCAAGTTCACAAATGTCAGAGCT 540
Qy 541 GGTGTTATCCATTTGGCTTAAGCTTGGCTTGTGAATGGGCTAAGTCAACATCAGAGT 600
Db 541 GGTGTTATCCATTTGGCTTAAGCTTGGCTTGTGAATGGGCTAAGTCAACATCAGAGT 600
Qy 601 AATCTTTAAACCCAGTTCATCTACGCTCTTTCAGCAAGATGTTATCATGTAAC 660
Db 601 AATCTTTAAACCCAGTTCATCTACGCTCTTTCAGCAAGATGTTATCATGTAAC 660
Qy 661 GAAAGATTTGACACAGATGATCTCTGATATCCCAACAAAGATGCCAACAAG 720
Db 661 GAAAGATTTGACACAGATGATCTCTGATATCCCAACAAAGATGCCAACAAG 720
Qy 721 GAATCATTTGCTGTTTGTACTGCTTCTGATGCTGCTTCACTACTGCT 780
Db 721 GAATCATTTGCTGTTTGTACTGCTTCTGATGCTGCTTCACTACTGCT 780
Qy 781 GCCAGCTTACTGTTGATGATGCTGTTCACTTCTGG 816
Db 781 GCCAGCTTACTGTTGATGATGCTGTTCACTTCTGG 816

RESULT 2
US-10-653-047-41
; Sequence 41, Application US/10653047
; Publication No. US20040229367A1
; GENERAL INFORMATION:
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; APPLICANT: Jeffrey R. Shuster
; APPLICANT: Sakari Kauppinen
; APPLICANT: Ib Groch Clausen
; APPLICANT: Peter Bjørke Olsen
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 5849.200-US
; CURRENT APPLICATION NUMBER: US/10/653,047
; PRIOR FILING DATE: 2003-08-29
; PRIOR APPLICATION NUMBER: US/09/533,559
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/273,623
; PRIOR FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 7860
```

```
SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 3203
; TYPE: DNA
; ORGANISM: Fusarium venenatum
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(3203)
; OTHER INFORMATION: n = A,T,C or G
US-10-653-047-41

Query Match 12.2%; Score 99.4; DB 20; Length 3203;
Best Local Similarity 52.7%; Pred. No. 4.5e-18;
Matches 265; Conservative 0; Mismatches 211; Indels 7; Gaps 2;

Qy 308 TGGTTACACAGAGCTGTTACTGTGAAGAACTTCCATGGAAGATTAACCAAGACG 367
Db 2614 TAGTACAGTGGCTGCTTACAGAGAAATTCAGAGCAATCACTATCCATGATGTA 2673
Qy 368 CTGAGAGATGATGAGGTTAACTTGTGGTCTTGTGATGTTTCAAGCTTGTCTA 427
Db 2674 TGGCAAGTTGTGGGCTGTTAATGTGACGTAATCTTGTGCAAGTTCAGTTC 2733
Qy 428 AGCATTTGATCAAGAGATGATCAAGGTCCTCTGTTGATGTTGATGTTATGTCG 487
Db 2734 A--ACATCTCATGAGACGCAAGTACCTGTGATGATGATGATGATGATGATG 2790
Qy 488 GTGCTATTTGATCAAGATCTCAAAACCAAGTTCACAAATGTCACAAATGTCAGAG 547
Db 2791 GTGCTATTTGATCAAGATCTCAAAACCAAGTTCACAAATGTCACAAATGTCAGAG 547
Qy 548 TCCATTTGGCTTAAGCTTGGCTTGTGAATGGGCTAAGTCAACATCAAGTAACTT 607
Db 2851 GTCACTGCTGCTTCCCTCGCAGATGAGTGGCTCAAGTCAAGTCAAGTCAAGT 2910
Qy 608 TAAACCAAGTTCATCTACGCTCTTTCAGCAAGATGTTATCATGTAACGAGAT 667
Db 2911 TCTCTCTGCTTACATGTCATGCTCTTCAAGCAAGATGTTATCATGTAACGAGAT 2970
Qy 668 TGTACACAGATGATGATGCTGTTATCCCAACAAAGATGCCAACAAGATTAAC 727
Db 2971 TCGAGAGACCTGACATCTTATTTCTCAGAGACGATGAGTCACTCAAGATCTGA 3030
Qy 728 TTGCTGCTGTTTGTACTGCTTCTGATGCTGCTTCACTACTGCT 787
Db 3031 TGGGCTGCTG---GACCTTTTGTGCAAGATGCTTCTTACATGATGCTGGGCAATC 3086
Qy 788 TACTGCTGATGATGCTTCACT 810
Db 3087 TTAGAGTATGAGAGATATACT 3109

RESULT 3
US-10-653-047-4932
; Sequence 4932, Application US/10653047
; Publication No. US20040229367A1
; GENERAL INFORMATION:
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; APPLICANT: Jeffrey R. Shuster
; APPLICANT: Sakari Kauppinen
; APPLICANT: Ib Groch Clausen
; APPLICANT: Peter Bjørke Olsen
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 5849.200-US
; CURRENT APPLICATION NUMBER: US/10/653,047
; PRIOR FILING DATE: 2003-08-29
; PRIOR APPLICATION NUMBER: US/09/533,559
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/273,623
; PRIOR FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 7860
```

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 4932
LENGTH: 687
TYPE: DNA
ORGANISM: Aspergillus oryzae
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(687)
OTHER INFORMATION: n = A,T,C or G
US-10-653-047-4932

Query Match 10.9%; Score 88.6; DB 20; Length 687;
Best Local Similarity 49.8%; Pred. No. 2.9e-15;
Matches 279; Conservative 0; Mismatches 275; Indels 6; Gaps 2;

```

QY 196 TTGAAGAAGTTCAGAAAGTGGTTCATATGCGTGTGATTTCTGATTCGATCCGTT 255
DB 114 TTGAACAAGTCCCAAGCTCACCGCCACTGCGTGAATTTCCGACCTTAACCTCGTC 173
QY 256 CACAAAGTGTTCCTCAAGTTCCTAAGATTTTGTGAAGTTCCTTCACTTGTGAAC 315
DB 174 AACGATGCCCTCTCCGATTTATCTCAAGCAGCGCAAGAT---CGACAACTGGTCAAC 230
QY 316 ACAGTGTGTCTGTAAGAACTTCCCATGTGAAGATTAACCCAGCCAGAACGCTGAGAG 375
DB 231 TCCGCGGATTCAGCAAACTTCGATGCAATCTCTTACCTTCACGACCGTGTGCAAAAG 290
QY 376 ATGTGAAGTTCATCTGTGGTTCCTTGTATGTTTCTCAAGCTTGTCTAAGCATTTG 435
DB 221 CTTGGGGCGTTAATGTGATGGAACATACCTTTTCGCAACCGGTGTGCCAAGC---AC 347
QY 436 ATCAAGAAGGTATCAAGGGTCTTCTGTGTTTGTATGTTGTCTATGTCTGTGCCAT 495
DB 348 CTCAATGAGCGCAAGTTCGCGGCAAGATTTGATGATGTGATGATGTCTGTGTATC 407
QY 496 GTCAACGATCTCTCAAAACCAAGTGTCTACACATGTCCAGGCTGTGTATTCATTG 555
DB 408 GACAACGTCGCGCAAGCCCAAGCTCTTACCAAGCGCGCAAGCGCTGTGTCTCAACT 467
QY 556 GCTAAGACTTGGCTGTGTAATGGGCTAAGTACATCAAGATTAATCTTTAACAACA 615
DB 468 GCCGCTACTTCGCGCGGAATGGCCGGTCAACATTCGGGTAGCTGATGAGCCCT 527
QY 616 GGTATCATCTACGGTCTTTTGAACCAAGATTTATCAATGTGAAGAAATTTATCAAC 675
DB 528 CGATACATGCTTACTGACCTGACCGCAAGATTTTGAATGAGAACCCCGAATGGGGAC 587
QY 676 AGATGATCTCTGATATCCCAACAAGAAATGTCCGAACCAAGGAATACATTTGTCT 735
DB 588 AAGTGATCTGCTCATCCCAACCGCAAGATGGGTACTCCGAGGACCTGATGGTCCC 647
QY 736 GTTTGTACTGCTTTCTGA 755
DB 648 GTTACCTTTCTGCTCAGGA 667

```

RESULT 4

US-10-369-493-36723
Sequence 36723, Application US/10369493
Publication No. US20030233675A1

GENERAL INFORMATION:

APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.

APPLICANT: Chen, Xianfeng

TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

FILE REFERENCE: 38-10(52052)B

CURRENT FILING DATE: 2003-02-28

PRIOR APPLICATION NUMBER: US 60/360,039

NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 36723
LENGTH: 637
TYPE: DNA
ORGANISM: Aspergillus nidulans
US-10-369-493-36723

Query Match 10.0%; Score 81.8; DB 17; Length 637;
Best Local Similarity 55.0%; Pred. No. 2.8e-13;
Matches 183; Conservative 0; Mismatches 147; Indels 3; Gaps 1;

```

QY 304 CACTTGCTTACACAGCTGTCTACTGTGAAAATTCCTCCATGTGAAGATTAACCAACCAAG 363
DB 70 CATCTGTACGCTGTGCGGATTTACCAAACTTCGACGCCATCAGCTACCAATACAGC 129
QY 364 AACCTGGAAGATGTGAAGTTACTTGTGGTCTTTGTATGTTTCTCAAGCTTT 423
DB 130 CGTATGCAAGACTTGGGGGTGAACGTGACGCAATATCTGTTGCTACGGGCTG 189
QY 424 GCTAAGCATGTATCAAGAAAGTATCAAGGGTCTTGTGTTTGTATGTTGTATG 483
DB 190 GCAAGCACTTGAT---GAGGCGCAAGCGCCCGGAGATATGTATGATGTGATGATG 246
QY 484 TCTGTGCTATGTATCAAGATCTTCAAAACCAAGTTGTCTACATGTCCAGGCTGT 543
DB 247 TCTGTGCTATGTATCAAGATCTTCAAAACCAAGGCTTCAACGCGCCAGGCAAGC 306
QY 544 GTTATCAATTTGGCTTCAAGCTTGGCTGTGATGGGCTAAGTATCAATCAGATTAAT 603
DB 307 GTTCCCACTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 366
QY 604 TCTTAAACCAAGTTACATCATCGTCTTGG 636
DB 367 TGCATTAACCTCGATACATGCTGATCTGCTG 399

```

RESULT 5

US-10-369-493-26991
Sequence 26991, Application US/10369493
Publication No. US20030233675A1

GENERAL INFORMATION:

APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.

APPLICANT: Chen, Xianfeng

TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

FILE REFERENCE: 38-10(52052)B

CURRENT FILING DATE: 2003-02-28

PRIOR APPLICATION NUMBER: US 60/360,039

NUMBER OF SEQ ID NOS: 47374

SEQ ID NO 26991

LENGTH: 735

TYPE: DNA

ORGANISM: Neurospora crassa

US-10-369-493-26991

Query Match 9.9%; Score 80.6; DB 17; Length 735;
Best Local Similarity 61.8%; Pred. No. 7e-13;
Matches 128; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

```

QY 436 ATCAAGAAGTATCAAGGGTCTTGTGTTTGTATGTTTCTATGTCTGTGCTCAT 495
DB 364 ATGATGGCGCTGCAAAACCGGCTCATATCTTGTGCTGTGATGCTCGGACCATC 423
QY 496 GTCAACGATCTCTCAAAACCAAGTTGTCTCAACATGTCCAGGCTGTGATCATTTG 555
DB 424 GTCACTACCTCTGAGGACAGTGTGTCAACACCTTCAAGGCTGTGATCATTCAGCTC 483
QY 556 GCTAAGACTTGGCTGTGATGGCTAAGTACAAACATCAGATTAATCTTTAAACCA 615

```


ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_88858C.1
US-10-425-115-172450

Query Match 8.3% Score 72.4; DB 20; Length 775;
Best Local Similarity 56.2% Pred. No. 1.9e-10;
Matches 136; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 383 AGGTAACTTGTGGGTTCTTTGTATGTTTCTCAAGCCCTTGTCAAGCCATGTATCAAG 442
DB 465 AGTCAAGGTTTGGGAGACGTTGCGCTGTCTCAAGCCGCTGCCACCTCTGACCAAGT 524
QY 443 AAGGTATCAAGGCTGCTTCTGTGTTTGTATGTTTCAATGTCTGAGTCAATGTCACG 502
DB 525 CTGGCTTCAAGACCGCTTCTCATGCTCTCACTCTGCTCAATGTTCTCCAAATGTCACAC 584
QY 503 ATCCCAAAACCAAGTTGTCTCAACATGTCACAGGCTGTGTATCCATTTGGCTAAGA 562
DB 585 GTGGCATCCACCAAGTCTTCTCAACCTGCTCAAGGCTGCGCTGTGATGTTAAGC 644
QY 563 CTTTGGCTTGTGAATGGGCTAAGTACACATCAGAGTTATCTTTAAACCAAGTTACA 622
DB 645 AGCTGCTGTGATGAGTGGCCGACAAAGGATTCAGTCAAGCTCTTGTCCAGGTTACG 704
QY 623 TC 624
DB 705 TC 706

RESULT 9
US-10-369-493-25810
Sequence 25810, Application US/10369493
Publication No. US2003023675A1

GENERAL INFORMATION:

APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xiandeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 473/74
SEQ ID NO 25810
LENGTH: 783
TYPE: DNA
ORGANISM: Schizosaccharomyces pombe
US-10-369-493-25810

Query Match 8.3% Score 67.6; DB 17; Length 783;
Best Local Similarity 51.2% Pred. No. 5.1e-09;
Matches 213; Conservative 0; Mismatches 194; Indels 9; Gaps 2;

QY 171 CAATACGCTACTGAAGATTAAGTTGAAGAAAGTTCCAAAGATGGGTTCAATGCTG 230
DB 171 CGAAAGGCTCAAGAAATGGCCAGCTAATGTGTCAAGACTTATACATGCAATG 230
QY 231 TGATATTTCTGATTTCTGATACCGTTACAGAGTGTGCTCAAGTTGCTAAGGATTTGG 290
DB 231 CGATCTACTATTTCTTAAAGAGTGAACATGCTTGTCTGAGATTCAAAGGCTTTGA 290
QY 291 TAAAGTGCATGCACTTGTGAACAGCTGTGTACTGTGAAAACCTTCCATGTGAAGA 350
DB 291 TACTATGATATGCTGTGCGCAACAGCGATATTGCACTGGCAAGTCCGCAAT---GA 347
QY 351 TTACCCAGCAAGACGTGAGAGATGTGAAGTTAACTTTGGGTTCTTTGTATGT 410
DB 348 CATGACTTACGAAGATTGTGTAATGAATAAAGTCAACTTGTCTGGTGTCTTCAAGCT 407

QY 411 TTTCAGCCTTTGCTAAGCATTTGATCAAGAGATATCAAGGTTCTTCTGTTT 470
DB 408 CGCTCACAATGCTGGCCCATCTTCCAAAGCAAG-----CATGCGAGCTTGTAGC 461

QY 471 GATGTTCTATATGTTGTGTCATTTGTCATGCAAGATCTTCAAAACCAAGTTGTATCAACT 530
DB 462 TACGCTTCAATGTTGTGTTGTTGTTTAAAGGTTCTTCAACCAAGATGTGCTTACAAAC 521

QY 531 GTCAAGGCTGTGTTATTCATTTGGCTAAGACTTGGCTTGTGAATGGGCTAAGT 586
DB 522 TTCAGGCGGTGTCATTTCACTATCATCAAGATTTGGCAATGCGCAAT 577

RESULT 10
US-10-369-493-26610
Sequence 26610, Application US/10369493
Publication No. US2003023675A1

GENERAL INFORMATION:

APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xiandeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 473/74
SEQ ID NO 26610
LENGTH: 774
TYPE: DNA
ORGANISM: Thermotoga maritima
US-10-369-493-26610

Query Match 7.8% Score 63.6; DB 17; Length 774;
Best Local Similarity 61.4% Pred. No. 7.7e-08;
Matches 102; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 472 ATGGTTCTATGCTGTGCTCATTTGTCACAGCTCTCAAAACCAAGTTGTCTACACATG 531
DB 436 ATCGGTCAATGCTCGGACACATGTCGAACAACTCAGAGCAGACGCTTACAGCCT 495
QY 532 TCAGAGCTGTGTATTCATTTGCTGCTAAGACTTGTGCTGTGAATGGGCTAAGTAC 591
DB 496 TCGAAAGCGGTGTGATCATCTCACAGATCTTGGCGCGGAGTGGGCGCCGATACGGA 555
QY 592 ATCAGAGTTATTTCTTAAACCCAGGTTACATCTACGCTCTTGA 637
DB 556 ATCAGGTTAACACATTAAGCCCGGATACATCAAGAACACCTCTCA 601

RESULT 11
US-10-653-047-3866
Sequence 3866, Application US/10653047
Publication No. US20040229367A1

GENERAL INFORMATION:

APPLICANT: Randy M. Berka
APPLICANT: Michael W. Rev
APPLICANT: Jeffrey R. Shuster
APPLICANT: Sakari Kauppinen
APPLICANT: Ib Groth Clausen
APPLICANT: Peter Bjarke Olsen
TITLE OF INVENTION: Methods for Monitoring Multiple Gene
FILE REFERENCE: 5849 200-US
CURRENT APPLICATION NUMBER: US/10/653,047
CURRENT FILING DATE: 2003-08-29
PRIOR APPLICATION NUMBER: US/09/533,559
PRIOR FILING DATE: 2000-03-22

QY 706 ATGTCGACCAAGGATATCATGTCGCTTTTGTACTGCTTCTGATCTGCT 765
DB 703 CGGCGTGGAAAGCCCAATGAGTGTCTGACTTAATAGCCCTTCTTGCTCCCTGCTGCT 762
QY 766 TCATCACTACTGTCGTCAGCTTACTGTTGATGTCGTTGACTTCT 813
DB 763 TCTTATATTAATCTGCGCAATTAATATGCGCTGATGTCGATTCACAGCT 810

RESULT 14

US-09-938-842A-948
; Sequence 948, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 948
; LENGTH: 789
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-948

Query Match 6.9%; Score 56.2; DB 9; Length 789;
Best Local Similarity 55.2%; Pred. No. 1.2e-05;
Matches 153; Conservative 0; Mismatches 118; Indels 6; Gaps 2;

QY 533 CCAAGGCTGCTGTTATCCATTGGCTTAAGACTTTGGCTTGAATGGGCTAATCAACA 592
DB 485 CAAAAGAGCTTTGATTCAGCTAGCTAAATAATTTGGCAATGTAAGGCAAAAGCGCA 544
QY 593 TCAGAGTTAATCTTTAAACCAAGTTACATCAAGCTTGTGACCAAGATGTTATCA 652
DB 545 TAAAGGCCAAGCCTGTGGCTTAATGATCAATATCTCTGTCTCAATCTTATCTTG 604
QY 653 ATGTAACGAAGATTGTAACAAGATGATCTTGTATCCCAACAAGAAATGTCG 712
DB 605 AGG--ACGTCACTTCAAGAGGCAATGTTAGTAGACTCACTGTCGTGTTGAG 661
QY 713 AACCAAGGAATATCATGTCGCTGTTTGTACTGTTCTGAAATGCTGCTTATACA 772
DB 662 AGCCAAATGAAG--TTGATCATCAATGAGCCCTTCTGTGCTTACTGCAAGCTTCTTATA 718
QY 773 CTACTGTGCGCAGCTTACTGTTGATGTCGTTTCTAC 809
DB 719 TTACTGTGTCAGCTATTGTTGTGATGAGGTCCTCAC 755

RESULT 15

US-09-938-842A-948
; Sequence 948, Application US/09938842A
; Patent No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPT300-3

; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 948
; LENGTH: 789
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-948

Query Match 6.9%; Score 56.2; DB 11; Length 789;
Best Local Similarity 55.2%; Pred. No. 1.2e-05;
Matches 153; Conservative 0; Mismatches 118; Indels 6; Gaps 2;

QY 533 CCAAGGCTGCTGTTATCCATTGGCTTAAGACTTTGGCTTGAATGGGCTAATCAACA 592
DB 485 CAAAAGAGCTTTGATTCAGCTAGCTAAATAATTTGGCAATGTAAGGCAAAAGCGCA 544
QY 593 TCAGAGTTAATCTTTAAACCAAGTTACATCAAGCTTGTGACCAAGATGTTATCA 652
DB 545 TAAAGGCCAAGCCTGTGGCTTAATGATCAATATCTCTGTCTCAATCTTATCTTG 604
QY 653 ATGTAACGAAGATTGTAACAAGATGATCTGTTATCCCAACAAGAAATGTCG 712
DB 605 AGG--ACGTCACTTCAAGAGGCAATGTTAGTAGACTCACTGTCGTGTTGAG 661
QY 713 AACCAAGGAATATCATGTCGCTGTTTGTACTGTTCTGAAATGCTGCTTATACA 772
DB 662 AGCCAAATGAAG--TTGATCATCAATGAGCCCTTCTGTGCTTACTGCAAGCTTCTTATA 718
QY 773 CTACTGTGCGCAGCTTACTGTTGATGTCGTTTCTAC 809
DB 719 TTACTGTGTCAGCTATTGTTGTGATGAGGTCCTCAC 755

Search completed: July 12, 2005, 15:54:44
Job time : 640 secs

This Page Blank (usptg)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2005, 11:47:19 ; Search time 3387 Seconds

(without alignment)
9170.488 Million cell updates/sec

Title: US-10-720-018-1

Perfect score: 816
Sequence: 1 atgactgactcattccaac.....atggtgttcactcttgg 816

Scoring table: IDENTITY NUC

Gapop 10'-0' , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hlc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	276.6	33.9	899	6	CNS077V0
2	129.2	15.8	969	9	AL433186.T7 end of
3	118.4	14.5	682	4	CD457882 FG04d.02e
4	113.4	13.9	713	7	BI750180 FG02.10b0
5	113.4	13.9	836	7	CF677471 CCA1590TF
6	113.4	13.9	836	7	CF701403 CCACT16TF
7	112.8	13.8	777	7	CF715948 CCA241TF
8	112.8	13.8	787	7	CF817896 EST695278
9	111.2	13.6	738	7	CF824542 EST701924
10	106.4	13.0	676	7	CF812382 EST689764
11	102.8	12.6	718	1	CF824543 EST701925
12	100.8	12.4	615	7	AU3638396
13	97.2	11.9	447	7	CO136295 EST830966
14	93.2	11.4	511	1	CO141255 EST835926
15	90.8	11.1	892	7	AU249728 AU249728
16	82.8	10.1	758	7	CN811863 FG09.08p2
17	77.4	9.5	753	5	CN812466 FG11.01e0
18	77.2	9.5	896	6	BQ855471 QGB26L1E
19	76.6	9.4	574	6	CF820440 EST697822
20	76	9.3	833	7	CD456321 FG03.03g0
21	76	9.3	839	7	CO012222 EST800561
22	76	9.3	922	7	CO012227 EST800562
23	76	9.3	949	7	CO009933 EST798268
24	76	9.3	963	7	CO004456 EST792791
					CO004455 EST792790

25	75.8	9.3	692	5	BQ992634
26	74.4	9.1	809	6	CB905184
27	74.4	9.1	809	7	CF876511
28	73.8	9.0	479	7	CF191935
29	73	8.9	816	2	BE642336
30	72.6	8.9	688	5	BP098973
31	72.6	8.9	735	7	CO140332
32	71.4	8.8	912	7	CO009934
33	70.8	8.7	690	7	CK569564
34	70.8	8.7	1048	9	CNS07622
35	70.2	8.6	712	6	CD458386
36	70	8.6	482	7	CF190316
37	68	8.3	734	7	CF447731
38	67.4	8.3	659	9	AJ638576
39	67.4	8.3	971	9	CNS07CVV
40	66.8	8.2	800	7	CO006699
41	65.4	8.0	786	5	BQ855467
42	65.2	8.0	868	7	CO006700
43	64.6	7.9	700	7	CK567406
44	63.8	7.8	621	5	BU014570
45	63.6	7.8	703	5	BQ861127

ALIGNMENTS

RESULT 1
LOCUS CNS077V0 899 bp DNA linear GSS 08-JUL-2001
DEFINITION T7 end of clone BB0AA014G1 of library BB0A from strain CBS 4732
ACCESSION AL433186
VERSION AL433186
KEYWORDS
SOURCE
ORGANISM
Pichia angusta
Bukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Pichia.

REFERENCE
AUTHORS Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G., Bolotin-Fukuhara,M., Bon,E., Broctier,P., Casaregola,S., de-Montigny,J., Dujon,B., Durren,P., Lepingle,A., Llorente,B., Malpertuy,A., Neuvéglise,C., Ozler-Kalogeropoulos,O., Potier,S., Saurin,M., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M., Wincker,P. and Weissenbach,J.
Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies
FEBS Lett. 487 (1), 3-12 (2000)

TITLE
JOURNAL MEDLINE
PUBMED
REFERENCE
AUTHORS Blandin,G., Llorente,B., Malpertuy,A., Wincker,P., Artiguenave,F. and Dujon,B.
Genomic exploration of the hemiascomycetous yeasts: 13. Pichia angusta
FEBS Lett. 487 (1), 76-81 (2000)

TITLE
JOURNAL MEDLINE
PUBMED
REFERENCE
AUTHORS
JOURNAL
COMMENT
Direct Submission
Submitted (08-SEP-2000) Genoscope - Centre National de Séquençage, 2 rue Gaston Crémieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail: secref@genoscope.cns.fr; Web: www.genoscope.cns.fr)
This GSS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces exiguus, Saccharomyces servazii, Zygosaccharomyces rouxii, Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of

the other extremity of this insert.
Location/Qualifiers
1..899
/organism="Pichia angusta"
/mol_type="genomic DNA"
/strain="CBS 4732"
/db_xref="taxon:4905"
/clone="BB0A014G11"
/clone_lib="BB0AA"
/note="end : T7"
misc_feature
18..633
/note="similar to P50167 [D-arabinitol 2-dehydrogenase
ARDH] [Pichia stipitis]
1 putative fructose 1,6
/evidence=not experimental
complement(479..812)
/note="similar to Saccharomyces cerevisiae ORF YKL126w [YPK1 ; ser/thr-specific protein kinase]"
/evidence=not experimental
complement(4732..812)
misc_feature
/note="similar to Saccharomyces cerevisiae ORF YMR104c [YPK2 ; ser/thr protein kinase]"
/evidence=not experimental

ORIGIN
Query Match 33.9%; Score 276.6; DB 9; Length 899;
Best Local Similarity 66.0%; Pred. No. 3.6e-68;
Matches 416; Conservative 0; Mismatches 210; Indels 4; Gaps 1;

QY 191 TGAAGTTGAAGAAAGTTCCAAAGATGGGTTCAATATGCTGATATTTCTGATTTCTGATA 250
DB 4 TCAAGAGAGATCTGTCGCAAGATGCACTGTAAGTTGGACATTAAGCAGCGTCGAGC 63
QY 251 CCGTTCAAGAGTGTTCCTCAAGTTCGAAGATTTTGTGAAGTTCGCAATTCGACTGG 310
DB 64 ACGTGAAGAGGTCCTTGTCTCAATATAGCGCGACTTTGAAAGTAAACCGGTCGATCG 123
QY 311 TTAACAGAGCTGTTACTGTGAAACTTCCATGTAAGATTAACCAAGCAAGAGCGTGG 370
DB 124 TCAACACGCTGTTATTCGAGAAATTCACGCGGACACATACCCCGCAAGAACGGCG 183
QY 371 AGAAGATGGTGAAGGTTAACTTGTGGGTTCTTGTATGTTTCTCAAGCGCTTGTAGC 430
DB 184 AAAAGCTTCTCAAGGTTCAACTTTTGGGGCGCTTGTAGCTGACAGACATTCGCAAC 243
QY 431 CATTGATCAAGAGATTCAGAGGTCCTTCTGTTTGTGATTTGATTTGATTCGATTCGG 490
DB 244 CACTTATCGATTAACAACATCAAGGAGGCTCATGTAATGATTCGATTCGAGAG 303
QY 491 CCATTGTCAAGATCTCTCAAAACCAAGTTGTCTACAAACAATGCCAAGGCTGGTTATTC 550
DB 304 AATGTCACAGACCAACCAAGCGCAAGTGCCTTCAACATGTCCAAGCGGCTGTGATTC 363
QY 551 ATTGGCTAAGACTTGGCTTGTGAATGGGCTAAGTCAACATCAAGTAAATTCCTTAA 610
DB 364 ACATGTCGAAGTCTTCGCGCGCTGAATGGGCAAAATACAGCATCAAGATGAATACCTGT 423
QY 611 ACCGAGTTACATTAACGTCCTTTGACCAAGATTTTCAATGTTAATGTTAAGAAATGT 670
DB 424 CGCGAGGATACATCTTACACCGCTCAAAAGACCTTATTAAGGTTAATTAAGAAATGT 483
QY 671 ACAACGATGATCTCTGGTATCCCAACAAGATGTCGAACCAAGAAATACATTTG 730
DB 484 ACCAGAGATGGCTTTGCTGACCCCAATGGGCGCTTTTCGGAAGTAAAGAAATTCACCG 543
QY 731 GTGCTGTTTGTACTTGTCTTGAATGCTGCTTCAATACATTAAGTGTCCAGCTTA- 789
DB 544 GTACGTTCTTAAGTCTCTCAATCTCGGCTTCTTATTACACCAAGGGAACAAATCAT 603
QY 790 ---CTGCTGATGATGTTCACTTCTGG 816
DB 604 TATTGTCTGACGTTGTTCACTTGTGG 633

RESULT 2
CD457882
LOCUS
DEFINITION
CD457882
R604d_02e10_A
Fg04d_02e10_A
Fusarium graminearum mycelium grown on wheat heads
Gibberella zeae cDNA clone Fg04d_02e10, mRNA sequence.
EST.
CD457882.1 GI:31372622
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Gibberella zeae
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
REFERENCE
Ouellet, R., Koul, A., Dan, H., Harris, L.J., Chapados, J., Couroux, P.,
De Moore, A., Hattori, J., Lacroix, C., Masotti, M., Robert, L.S.,
Singh, J.A., Sprott, D. and Tinker, N.A.
Fusarium graminearum mycelium grown on wheat heads under high
humidity conditions
Unpublished (2003)
CONTACT
Ouellet, Therese
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-Food Canada
Neelby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A 0C6,
CANADA
Tel: (613) 759-1658
Fax: (613) 759-1701
Email: ouellet@agr.gc.ca.

FEATURES
source
1..969
/organism="Gibberella zeae"
/mol_type="mRNA"
/strain="DAOM 180378"
/db_xref="taxon:5518"
/clone="Fg04d_02e10"
/tissue_type="mycelium"
/dev_stage="Asexual"
/lab_host="E. coli"
/clone_lib="Fg04_AAFc_ECORC_Fusarium_graminearum_mycelium_
grown on wheat heads"
/note="Vector: Bluescript SK+; Site 1: EcoRI; Site 2:
XhoI; Fusarium grown on wheat (cv. Roblin) under high
humidity. cDNA made using Stratagene kit."

ORIGIN
Query Match 15.8%; Score 129.2; DB 6; Length 969;
Best Local Similarity 57.1%; Pred. No. 1.1e-25;
Matches 289; Conservative 3; Mismatches 206; Indels 8; Gaps 3;

QY 305 ACTTGTTAACAAGCTGTTACTGTGAAAACCTTCCATGTAAGATTAACCAAGCAAGA 364
DB 254 ACGTGCACCTGACGCTGCTTCAAGGAACTTGAAGCGGTAAATCAACCCATCGACC 313
QY 365 ACGTGAAGAAGTGTGAAGTTAACTTGTGGTCTTGTATGTTCTCAAGCCTTGG 424
DB 314 GTTCTCGTAAGCTTGGCTGTAAAGTTGACGTAATCTCTTGCACATCAATAGTCG 373
QY 425 CTAAGCATTTGAT-CAAGAAGTATCAAGGTCCTTCTGTTGTTGATTGTTCTATG 483
DB 374 CGAGCACTTGAATGCAAGAAAGGCTCTGTGATGATC-----GTATGATTTGTAGCATG 429
QY 484 TCTGTGCAATTTGTAACGATCTCTCAAAACCAAGTTGTCTACACATGTCGAAGCTG 543
DB 430 TCGGATCATTTGTCAAGTTCTCTCASCCTCAAGGCTCTCTTAATGCGCCAAAGCCG 489
QY 544 GTTATCAATTTGGCTTAAGACTTGTGCTTGTGAATGGGCTAAGTAAATCAATCAAGATTAT 603
DB 490 GTGCGCATCTCGCTGCTTCTTCTTGGCCGTGAAATGGGCTCAAGCAACATCCGAGTCAAC 549
QY 604 TCTTTAAACCAAGGTTACATCAAGGTCCTTGAACCAAGATTTATCAATGTAACGA 663
DB 550 TGCATCTCTCCGGTTACATGTTGACTGACTCATCAAGAAATCTTGCAGCAACACCG 609

Oy	664	GAATGTCAACAGANGGATCTCTGGTATCCACAACAAAGATGTCGACCAAGAA	723
Db	610	GATTCAGAGCCAGATGGACTTCCCTTATCCCCAGGGCAAAATGGGACACCGAGAC	669
Oy	724	TACATTGGTGGCTGTTTGTATCTGTGCTTCGATCTGCTTCATACCTACTGGTGGC	783
Db	670	CTCATGGGGCCCGTGGCATTCCTCTTATGAG--TGCTTTCGTATGTACTGGTGGC	726
Oy	784	AGCTTACTGTTGATGATGTTTCAC	809
Db	727	CACATCCGASTCGATGGCGSCTACAC	752

RESULT 3	682 bp	mRNA	linear	EST 14-JUN-2004
BI750180	BI750180.1	GI:15771982		
LOCUS	BI750180			
DEFINITION	Fg02_10b03_R Fg02_AAFC_ECONC_Fusarium_graminearum mycelium Gdbberella zeae cDNA clone Fg02_10b03, mRNA sequence.			
ACCESSION	BI750180			
VERSION	BI750180.1			
KEYWORDS	EST.			
SOURCE	Gibberella zeae			
ORGANISM	Gibberella zeae			

REFERENCE

AUTHOR

TITLE	Expressed Sequence Tags from <i>Fusarium graminearum</i> mycelium
JOURNAL	Unpublished (2001)
COMMENT	Contact: Harris, Linda J.

FEATURES
source 1. .682

```

/organism="Gibberella zeae"
/mol_type="mRNA"
/strain="DAOM 180378"
/db_xref="taxon:5518"
/clone="Fg02_10b03"
/tissue_type="Mycelial tissue"
/dev_stage="Asexual"
/lab_host="E. coli (Sure cells)"
/clone_lid="Fg02 APC ECORC Fusarium graminearum mycelium"
/note="Vector: Bluescript SK+/XhoI-ECOR1; Site_1: EcoRI; Site_2: XhoI; Mycelial tissue was collected from V8 agar plates after a growth period of 6-7 days at 25 C with 14 hrs (P/L/UV) day light exposure. Mycelia was ground in liquid nitrogen prior it's storage at -80 C until RNA extraction. Directional cloning with 5' end of cDNA cloned into EcoRI site of pBluescript and 3' end of cDNA cloned into XhoI site of pBluescript (Stratagene, La Jolla, CA)."

```

Query March Similarity	14.5%;	Score 118.4;	DB 4;	Length 682;
Best Local Similarity	52.1%;	Pred. No. 1.2e-22;		
Matches 342; Conservative	8;	Mismatches 295;	Indels 11;	Gaps 4;

Qy 14 TTCACCTTTTGGATTCAGTGGCACTTACACATGTGACAGGGGCGCGGGTGGTTAG 73

Db 20 TGGCCACCTTTACCTCAGCGGTCAATGTTGGCATGTGACCGGTGTGCACGTGGGTTGG 79

Qy 74 CTGACCTTAACTCAAGGATTGTTGGGCTTACGGTTCTTGACATTGCTTGGTTGATATCG 133

Db 80 GTCTTGTTATGGGTCAAGATAGTAAATCTCTGGATTCACACTTGGCTCTTGTTGATATTTW 139

Oy	134	ACCAAGAAAAGCTCTGCTCAAAACAAGCCGAATACCAAAAATA---CGCTACTGAGAAT	190
Db	140	ATAAGAAAGAAAGCAGAAAGCAGACTAGCTTATATCTTGAAGAGTTCAAAAGAGAACCC	199
Oy	191	TGAAGTTGAAAGAAAGTTCCAAAGATGGGTTCAATATGCCCTGATATTTCTGATTCGATA	250
Db	200	CTCGAGCCCGAGATCCCAAAAGTCACTGGCCATTATGCTGATGTATGTATTCCTGAAAT	259
Oy	251	CCGTTCAACAAGGTGTTTGCTCAAGTTGCTAAAGATTTTGGTAAGTTGCCATTGCACCTGG	310
Db	260	CTGTGAGGCTTGTTTAAAGGTTTAAAGAGACGAGAAAGATGCACA---ACCTGG	316
Oy	311	TTAACACAGCTGCTTACTGTGAAAACTTCCCATGTGAAGATTACCCAGCAGAAAGCTGG	370
Db	317	TCACCTCAGCTGGCTTCAACGAGAACTTCGAAGCCGTTAACTACBSACATCGACCGTCTCC	376
Oy	371	AGAAGATGTAAGGTTAATTGTTGGGTTCTTGTATGTTCTCGAAGCCTTGCTGAAGC	430
Db	377	GTAAGCTTTGGGCTGTTAAGCTTGACGGTACATATTTCTTGGCAACATCAGTCGCGCAGGY	436
Oy	431	CATTGAT--CAAGAAGGTATCAAGGGTCTTCTGTGTTTGAATGGTTCTAATGTCTGGT	489
Db	437	ACTTTATGCAAGAAAGGCTCTCGTAGATC---GTCAATATTTTATGATGTATTCGGA	492
Oy	490	GCCATGTGCAAGATCTCTCAAAACCAAGTTGTCTAACCATGTGCCAAGCTGGTATTC	549
Db	493	TCCATATGTCAACGTTCTCAGACCTCAGGCTCCATATATATGCGCCAAAGCTGTGTGGCC	552
Oy	550	CATTGGCTAAGACTTTGGCTTTGGATGGGCTAAGTACAACATGACAGTTAATTTCTTGA	609
Db	553	CAACTCGGTCTTCTTTGGCCGTGTGAATGGGCTCAGGCAAACTCCGAGTTAATGTCAATC	612
Oy	610	AACCCAGGTTACATCTACGGTCTTTGACGAAGAATGTATCAATGATGAAGAGAGA	665
Db	613	TCTCCGGTTACATGTTGACTGCACTCACTCAGAAAGAACTTTGACAGCAACCCBGA	668

RESULT 4				
CF677471/c				
LOCUS				
DEFINITION	713 bp	RNA	linear	EST 16-AUG-2000
ACCESION	CF677471	CCAI990TF	C. neoformans strain JRG21	Cryptococcus neoformans var.
VERSION	CF677471.1	neoformans	CDNA clone CCAI990,	RNA sequence.
KEYWORDS	CF677471.1	GI:41531630		
SOURCE	EST.			
	Cryptococcus neoformans	var. neoformans	(Filobasidiella)	neoformans

ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
<p><i>Cryptococcus neoformans</i> var. <i>neoformans</i> Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae <i>Tremellasterella</i>.</p>	<p>1 (bases 1 to 713)</p>	<p>Loftus, B.</p>	<p>End sequencing of clones from a full length enriched, normalized JEC21 cDNA library</p>	<p>Unpublished (2003)</p>	<p>Other ESTs: CCA1990TR</p>

Contact: Brendan Loftus
 TIGR
 Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
 Rockville, MD 20850, USA
 Tel: 301-838-3543
 Fax: 301-838-0208
 Email: crypt@tigr.org
 Seq primer: TF.
 Location/Qualifiers
 1..713

```

FEATURES
    source
        Location/Qualifiers
            1..713
                /organism="Cryptococcus neoformans var. neoformans"
                /mol_type="mRNA"
                /strain="JEC21"
                /db_xref="taxon:40410"
                /clone="CAI190"
                /clone_lib="C.neoformans strain JEC21"

```

/note="Vector: pCMVSPORT6; Site 1: NotI EcoRV; The full length, normalized library was prepared from a variety of conditions using RNA provided by Joseph Heitman and Jennifer Lodge"

ORIGIN

Query Match 13.9%; Score 113.4; DB 7; Length 713;
 Best Local Similarity 54.6%; Pred. No. 3.5e-21; Mismatches 216; Indels 15; Gaps 2;
 Matches 278; Conservative 0; Mismatch 216; Indels 15; Gaps 2;

308 TGGTTACACAGCTGTTACTGTGTAAGAACTTCCCATGTGAAGATTACCAAGCCAGAACG 367
 649 TCGTCACTGCTGCGGATTTGTCGAAACTTTGTGCTCAAGATACCCATCATAGAGA 590

368 CTGAGAAAGATGTGAAGATTACTTGTGGTCTTTGTATGTTTCTCAAGCTTGTCTA 427
 589 TCAGAAAGCTGTGACATCAACATTATGGTACTGATTTGGCATTGAAGCTGCCA 530

428 AGCCATTGATCAAGAAAGATTCAGAGGTGCTTGTGTTGTTGATTGGTTCTATGCTG 487
 529 AGCTTATGCTGAA-----GGTGGTTCATTACCTTGTGCTCATTTATGAGCG 482

488 GTGCAATTGTCAAGATCTCTCAAAACCAAGTTGTCTACAGATGTCCAAAGGCTGTGTTA 547
 481 GTAGCATTTGTCAAGCTTCTCAACTCAACCCCTTACACTTTTCCAAAGCTGTGTC 422

548 TCCATTGTGGCTAAGACTTTGGCTGTGTGAATGGCTTAAGTACATCAAGATTATTTT 607
 421 GACACATGGCTGATCCCTGCGCGTGAATGGGCTCTCAAGGGTATCCGTGTCAACGCTC 362

608 TAAACCCAGGTTACATCTAGAGTCTTTGACCAAGAAATGTTATCAATGATACGAAGAT 667
 361 TTAGTCCCGGTTACGCTCTCAACCACTTGACTTAAGGTCAATCTGACGCCCAACCCGCTTC 302

668 TGTACACAGATGATCTCTGGTATCCCAACAAGATGTCCGAACCAAGAAATACA 727
 301 TCCGTAAGAGATGCTCAACCTATCCCATGGGTGATGGCCGACCTTCTGATCTCA 242

728 TTGGTGTGTTTGTACTTGTCTTCTGAATCTGCTGCTTCAATCACTAGTGTGCACT 787
 241 AGGGTCCGTCATTATCACTTGTCTGGA---CAGCTCAAGTACACCACTGTGTGAGA 185

788 TACTGTGATGATGTTTCACTTCTTG 816
 184 TCATGATTGACGGCGGTACACTTGTCTTG 156

RESULT 5
 CF701403/c 836 bp mRNA linear EST 16-AUG-2004
 LOCUS CCACCT16T.C.neoformans strain JEC21 Cryptococcus neoformans var.
 DEFINITION
 accession CF701403
 version CF701403.1 GI:41555562
 keywords EST.
 source Cryptococcus neoformans var. neoformans (Filobasidiella neoformans var. neoformans)
 ORGANISM
 Cryptococcus neoformans var. neoformans
 Eukaryota; Fungi; Basidiomycota; Hymenomycetes;
 Heterobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae;
 Filobasidiella.
 1 (bases 1 to 836)
 REFERENCE
 Loftus, B.
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Other_ESTs: CCAC16T
 Contact: Brendan Loftus
 TIGR
 Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
 Rockville, MD 20850, USA
 Tel: 301-838-3543
 Fax: 301-838-0208

Email: crypt@tigr.org
 Seq primer: TF.
 Location/Qualifiers
 1..836
 /organism="Cryptococcus neoformans var. neoformans"
 /mol_type="mRNA"
 /strain="JEC21"
 /db_xref="taxon:40410"
 /clone="CCAC16"
 /clone_1b="C.neoformans strain JEC21"
 /note="Vector: pCMVSPORT6; Site 1: NotI EcoRV; The full length, normalized library was prepared from a variety of conditions using RNA provided by Joseph Heitman and Jennifer Lodge"

ORIGIN

Query Match 13.9%; Score 113.4; DB 7; Length 836;
 Best Local Similarity 54.6%; Pred. No. 3.7e-21; Mismatches 216; Indels 15; Gaps 2;
 Matches 278; Conservative 0; Mismatch 216; Indels 15; Gaps 2;

308 TGGTTACACAGCTGTTACTGTGTAAGAACTTCCCATGTGAAGATTACCAAGCCAGAACG 367
 665 TCGTCACTGCTGCGGATTTGTCGAAACTTTGTGCTCAAGATACCCATCATAGAGA 606

368 CTGAGAAAGATGTGAAGATTACTTGTGGTCTTTGTATGTTTCTCAAGCTTGTCTA 427
 605 TCAGAAAGCTGTGACATCAACATTATGGTACTGATTTGGCATTGAAGCTGCCA 546

428 AGCCATTGATCAAGAAAGATTCAGAGGTGCTTGTGTTGTTGATTGGTTCTATGCTG 487
 545 AGCTTATGCTGAA-----GGTGGTTCATTACCTTGTGCTCATTTATGAGCG 498

488 GTGCAATTGTCAAGATCTCTCAAAACCAAGTTGTCTACAGATGTCCAAAGGCTGTGTTA 547
 497 GTAGCATTTGTCAAGCTTCTCAACTCAACCCCTTACACTTTTCCAAAGCTGTGTC 438

548 TCCATTGTGGCTAAGACTTTGGCTGTGTGAATGGCTTAAGTACATCAAGATTATTTT 607
 437 GACACATGGCTGATCCCTGCGCGTGAATGGGCTCTCAAGGGTATCCGTGTCAACGCTC 378

608 TAAACCCAGGTTACATCTAGAGTCTTTGACCAAGAAATGTTATCAATGATACGAAGAT 667
 377 TTAGTCCCGGTTACGCTCTCAACCACTTGACTTAAGGTCAATCTGACGCCCAACCCGCTTC 318

668 TGTACACAGATGATCTCTGGTATCCCAACAAGATGTCCGAACCAAGAAATACA 727
 317 TCCGTAAGAGATGCTCAACCTATCCCATGGGTGATGGCCGACCTTCTGATCTCA 258

728 TTGGTGTGTTTGTACTTGTCTTCTGAATCTGCTGCTTCAATCACTAGTGTGCACT 787
 257 AGGGTCCGTCATTATCACTTGTCTGACAGCTCCAG---TACACCACTGTGTGAGA 201

788 TACTGTGATGATGTTTCACTTCTTG 816
 200 TCATGATTGACGGCGGTACACTTGTCTTG 172

RESULT 6
 CF715948/c 836 bp mRNA linear EST 16-AUG-2004
 LOCUS CCAB241T.C.neoformans strain JEC21 Cryptococcus neoformans var.
 DEFINITION
 accession CF715948
 version CF715948
 keywords EST.
 source Cryptococcus neoformans var. neoformans (Filobasidiella neoformans var. neoformans)
 ORGANISM
 Cryptococcus neoformans var. neoformans
 Eukaryota; Fungi; Basidiomycota; Hymenomycetes;
 Heterobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae;
 Filobasidiella.
 1 (bases 1 to 836)
 REFERENCE
 Loftus, B.
 AUTHORS

TITLE End sequencing of clones from a full length enriched, normalized JEC21 cDNA library
JOURNAL Unpublished (2003)
COMMENT Other ESTs: CCAB241TR
 Contact: Brendan Loftus
 TIGR
 Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-3543
 Fax: 301-838-0208
 Email: crypt@tigr.org
FEATURES
 source
 Location/Qualifiers
 1..836
 /organism="Cryptococcus neoformans var. neoformans"
 /mol_type="mRNA"
 /strain="JEC21"
 /db_xref="taxon:40410"
 /clone="CCAB241"
 /clone_1lb="C.neoformans strain JEC21"
 /note="Vector: pCMVSPORT6; Site 1: NotI EcoRV; The full length, normalized library was prepared from a variety of conditions using RNA provided by Joseph Heitman and Jennifer Lodge"

ORIGIN
 Query Match 13.9%; Score 113.4; DB 7; Length 836;
 Best Local Similarity 54.6%; Pred. No. 3.7e-21;
 Matches 278; Conservative 0; Mismatches 216; Indels 15; Gaps 2;

QY 308 TCGTTTAAACAGCTGTTACTGTAAGAACTTCCATGTAAGATTACCAAGCAAGACG 367
 DB 670 TCGTCACTGCTCCCGGTATGTCGAAACTTGTGCTCAAGATGCCATCGATTAAGA 611
 QY 368 CTGAGAAGATGTAAGGTTAACTTGTGGGTTCTTGTATGTTTCTCAAGCCTTGTCTA 427
 DB 610 TCAAGAGCTGTGACATCAATATGAGTACTGTATGCGCACTTGAAGGCTGCCA 551
 QY 428 AGCCATTGATCAAGAGATATCAAGGGTCTTCTGTGTTTGTGATGTTCTATGTCG 487
 DB 550 AGCTTATGCTGAA-----GGTGTTCATTACCTCGCGGATCTATGAGCG 503
 QY 488 GTGCCATTGTCAAGATCTCTCAAAACAAAGTTGTCTACCAACATGTCGAAGGCTGTGTA 547
 DB 502 GTAGCATGTCAACCTTCTCTCAACCTCAAAACCTTACCACTTTCGAAGGCTGTGTC 443
 QY 548 TCCATTGGCTAAGCTTGTGGTGTGATGAGGCTAAGTCAACATCAAGTAAATCTT 607
 DB 442 GACACATGCTGATCTCGCGGTGGAATGAGGCTCTCAAGGTTATCCGTGCAACGCTC 383
 QY 608 TAAACCCAGTTACATCTACGGTCTCTTGACCAAGATGTTATCAATGATGTAAGAAAT 667
 DB 382 TTAGTCCGGTTACTCTCTCAACCACTTGAAGTCACTTCTCAAGCGCAACCCGTTTC 323
 QY 668 TGTACACAGATGATCTCTGTGATCCCAACAAAGATGTCCGAACCAAGAAATCA 727
 DB 322 TCCGTAAGAGTGGCTCAACCGTATCCCATGAGTGGCGGACCTTGTGATCTCA 263
 QY 728 TTGGTGTGTTTGTATCTTCTTGTGATCTGTGCTTCAATCACTACTGTGCGACGT 787
 DB 262 AGGGTCCGTCATTTACCTTGTCTGACAGCTTCAG--TACACACTGTGTGTGAGA 206
 QY 788 TACTGTGATGATGTTTCACTTCTTGG 816
 DB 205 TCAATATGACGCGGTTACACTTCTTGG 177

RESULT 7
 CF817896 777 bp mRNA linear EST 01-APR-2004
 LOCUS CF817896/6
 DEFINITION EST695278 Coccidioides posadasii saprobic phase cDNA library, greater than 4kb Coccidioides posadasii cDNA clone CIBB147 5' end, mRNA sequence.

ACCESSION CF817896
VERSION CF817896.1 GI:45923774
KEYWORDS EST.
SOURCE Coccidioides posadasii
ORGANISM Coccidioides posadasii
REFERENCE Eukaryotes; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Onykenales; mitosporic Onykenales; Coccidioides.
AUTHORS 1 (bases 1 to 777)
TITLE Analysis of gene expression in Coccidioides posadasii mycelia and spherules via expressed sequence tags
JOURNAL Unpublished (2003)
COMMENT Other ESTs: EST695277
 Contact: Gardner MJ
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301 838 3519
 Fax: 301 838 0208
 Email: gardner@tigr.org
FEATURES
 source
 Location/Qualifiers
 1..777
 /organism="Coccidioides posadasii"
 /mol_type="mRNA"
 /strain="C735"
 /db_xref="taxon:199306"
 /clone="CIBB147"
 /dev_stage="saprobic phase (mycelia)"
 /lab_host="E. coli DH10B; T1 phage resistant"
 /clone_1lb="Coccidioides posadasii saprobic phase cDNA library, greater than 4kb"
 /note="Vector: pXpress 1; Site 1: Not I; Site 2: Eco RV; Coccidioides posadasii saprobic phase cDNA library, size fractionated cDNA > 4 kb"

ORIGIN
 Query Match 13.8%; Score 112.8; DB 7; Length 777;
 Best Local Similarity 54.1%; Pred. No. 5.3e-21;
 Matches 275; Conservative 0; Mismatches 227; Indels 6; Gaps 2;

QY 304 CACTGTGTTAACAAGCTGTGTTACTGTAAGAACTTCCATGTAAGATTACCAAGCAAG 363
 DB 708 CATGTGTGTCACATCTGCGGGTTCACCTGAACCTTGAAGCGGTCTCCATCCGATAT 649
 QY 364 AACGTGAAGATGTAAGGTTAACTTGTGGTCTTGTATGTTTCTCAAGCCTT 423
 DB 648 CCGATGCAAGAACTGTGCTGTCACAGTGAAGGATATCTTTGCAAGTCTGTT 589
 QY 424 GCTAAGCCATTGATCAAGAGTATCAAGGGTCTCTGTGTTTGTGATGTTCTATG 483
 DB 588 GCGAAACATGAT---GGCTGCAAGTGGCTGACACATTTGTTTCAATGGAAGATG 532
 QY 484 TCTGTGCTATTGTCAAGATCTCAAAACCAAGTTGTCTAACAATGTCCAGGCTGCT 543
 DB 531 TCTGTGCTCAATGTATAGTCTCAGCCAGCAAGGCGGTCAATGTGCAAGGCTGCG 472
 QY 544 GTTATTCATTTGGCTAAGACTTGTGGCTGTGAATGGGTTAAGTACAACTAGATTAAT 603
 DB 471 AATAGGCACTTGGCTCTCTTGTGCAAGTAAATGGGCAAGTGGGAATCAAGATGAAC 412
 QY 604 TCTTAAACCAAGTTACATCTACGGTCTTGTGACCAAGATGTTATCAATGTTAAGCA 663
 DB 411 TGTATCACTCAAGATATCATTTAGTCTGCTGACCGCAAGATTTCTTGACGACACCT 352
 QY 664 GAATGTACAACAGATGATCTGTGTATCCCAACAAAGATGTCCGAACCAAGAA 723
 DB 351 GATTCACAAGAGAGATGACCTCTCATTTCAACAGCAAAATGGGTACACCTGAAGAC 292
 QY 724 TACATGTGTGCTTTTGTACTGCTTCTGAAATCTGTGCTTCACTATCACTAGTGTGCC 783
 DB 291 CTCATGGCCCAAGTACCTTTATTAAGTGA---TGGAGCAAAATATGTTACTGGGGCA 235
 QY 784 AGCTTACTGTGATGATGTTGCTTCACTT 811

Db 234 GATTGAGGTTGATGGCGGTACCTT 207

RESULT 8
CF824542/c 787 bp mRNA linear EST 01-APR-2004
LOCUS EST701924 Coccidioides posadasii saprobic phase cDNA library, 2 to
DEFINITION 4 kb Coccidioides posadasii cDNA clone CIDAR90 3' end, mRNA
sequence.
ACCESSION CF824542.1 GI:45930599
VERSION CF824542.1
KEYWORDS EST.
SOURCE Coccidioides posadasii
ORGANISM Coccidioides posadasii
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Onygenales; Mitosporic Onygenales; Coccidioides.
REFERENCE 1 (bases 1 to 787)
Gardner, M.J. and Cole, G.T.
Analysis of gene expression in Coccidioides posadasii mycelia and
spherules via expressed sequence tags
Unpublished (2003)
JOURNAL Other ESTs: EST701925
COMMENT Contact: Gardner MJ
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@ligr.org

FEATURES
source location/Qualifiers
1..787
/organism="Coccidioides posadasii"
/mol_type="mRNA"
/strain="C735"
/db_xref="taxon:199306"
/clone="CIDAR90"
/dev_stage="saprobic phase (mycelia)"
/lab_host="E. coli DH10B, T1 phage resistant"
/clone_lib="Coccidioides posadasii saprobic phase cDNA
library, 2 to 4 kb"
/note="Vector: pExpress 1; Site_1: Not I; Site_2: Eco RV;
Coccidioides posadasii saprobic phase cDNA library, size
fractionated cDNA 2 to 4 kb"

ORIGIN

Query Match 13.8%; Score 112.8; DB 7; Length 787;
Best Local Similarity 54.1%; Pred. No. 5.3e-21;
Matches 275; Conservative 0; Mismatches 227; Indels 6; Gaps 2;

QY 304 CACTTGTTTAAACAGCTGTTACTGTGAAACTTCCCATGTGAAGATTACCCAGC 363
DB 706 CATCTGTGACACTGCGGGGTTCACTGAAACTTTGACGCCGTGTCCTACCGCATGAT 647
QY 364 AACGTGAGAGATGGTGAAGTTAATTGTTGGCTTTCTGTATGTTTCTCAAGCCTTT 423
DB 646 CGCATGCAAGAACTGCTGCTCCGTAACGTGATGGTCAATCTTTTGCAGTTCCTGTT 587
QY 424 GCTAAGCCATTGATCAAGAGATCAAGGGTCTGTTGTTTGTATGTTGTTCTATG 483
DB 586 GCGAAACATCTAT--GGCTGCAAGTCCCTGGCAGCATTTGTTCATGGAAGCATG 530
QY 484 TCTGTGTCATTGTCAACGATCTTCAAAACAAAGTTTCTTACAACATGTCGAAGCTGAT 543
DB 529 TCTGTGTCATTGTTAACGTGCTCAAGCAAGGCCCTTACAACTGTCGAAGGCTGG 470
QY 544 GTTATTCATTGGCTTGAAGCTTTGGCTTGTGAATGGGCTTAAGTACAACATCAAGATTAT 603
DB 469 ATTAAGGCACTTGCTGCTTCTTGGCAGTTGAATGGGCAAGTGGGAATCAGAGTGAAC 410
QY 604 TCTTTTAAACCCAGTTACATCTACGCTCTTTGACCAAGAAATGTTATCAATGTTAACGA 663
DB 409 TGTATTCATGTCAGGATATCATGTTGACTGCTGACCCGCAAGATTTCTTGACCAACCTCT 350

QY 664 GAATTTGACAAACAGATGATCTGTATCCACACAAGAAATGTCCTCCAAAGGAA 723
DB 349 GATCTCAAGAGAGATGAGACCTCCCTCATTCACAAAGCAAAATGGGTATCACTGAAGAC 290
QY 724 TACATTTGCTGCTTTTGTACTTCTTTCGAATCTGCTGTTATACATCACTGAGGCC 783
DB 289 CTCATGGGCCCAAGAGAGCTTTCTTATTAAGTGA--TGGAGCAAAATATTTACTGGGCA 233
QY 784 AGTTACTGTTGATGGTGGTTTCACTT 811
DB 232 GATTGAGGTTGATGGCGGTACCTT 205

RESULT 9
CF812382/c 738 bp mRNA linear EST 01-APR-2004
LOCUS EST689764 Coccidioides posadasii saprobic phase cDNA library,
DEFINITION greater than 4kb Coccidioides posadasii cDNA clone CIBA394 3' end,
mRNA sequence.
ACCESSION CF812382
VERSION CF812382.1 GI:45918260
KEYWORDS EST.
SOURCE Coccidioides posadasii
ORGANISM Coccidioides posadasii
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Onygenales; Mitosporic Onygenales; Coccidioides.
REFERENCE 1 (bases 1 to 738)
Gardner, M.J. and Cole, G.T.
Analysis of gene expression in Coccidioides posadasii mycelia and
spherules via expressed sequence tags
Unpublished (2003)
JOURNAL Other ESTs: EST689765
COMMENT Contact: Gardner MJ
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@ligr.org

FEATURES
source location/Qualifiers
1..738
/organism="Coccidioides posadasii"
/mol_type="mRNA"
/strain="C735"
/db_xref="taxon:199306"
/clone="CIBA394"
/dev_stage="saprobic phase (mycelia)"
/lab_host="E. coli DH10B, T1 phage resistant"
/clone_lib="Coccidioides posadasii saprobic phase cDNA
library, greater than 4kb"
/note="Vector: pExpress 1; Site_1: Not I; Site_2: Eco RV;
Coccidioides posadasii saprobic phase cDNA library, size
fractionated cDNA > 4 kb"

ORIGIN

Query Match 13.6%; Score 111.2; DB 7; Length 738;
Best Local Similarity 53.9%; Pred. No. 1.5e-20;
Matches 274; Conservative 0; Mismatches 228; Indels 6; Gaps 2;

QY 304 CACTTGTTTAAACAGCTGTTACTGTGAAACTTCCCATGTGAAGATTACCCAGC 363
DB 612 CATCTGTGACACTGCGGGGTTCACTGAAACTTTGACGCCGTGTCCTACCGCATAT 553
QY 364 AACGTGAGAGATGGTGAAGTTAATTGTTGGCTTTCTGTATGTTTCTCAAGCCTTT 423
DB 552 CGCATGCAAGAACTGCTGCTCCGTAACGTGATGGTCAATCTTTTGCAGTTCCTGTT 493
QY 424 GCTAAGCCATTGATCAAGAGATCAAGGGTCTTGTGTTTGTATGTTGTTCTATG 483
DB 492 GCGAAACATCTGAT--GGCTGCAAGTCCCTGGCAGCATTTGTTCATGGAAGCATG 436
QY 484 TCTGTGTCATTGTCAACGATCTTCAAAACAAAGTTTCTTACAACATGTCGAAGCTGAT 543
DB 435 TCTGTGTCATTGTTAAGTGTCTCAAGCAAGGCCGTACCAATGTCGAAGAGTTGGC 376

OY	GTATCCATTGGCTAAAGACTTTGGCTTGTGAATGGGCTAAAGTCAAACTTCAAGATTAT	603
Db	ATAAAGCACTTGGCTGCTTCCCTGGCAATTGAATGGGCAAGTGTGGGAATCAGATGAAAC	316
OY	TCTTTAAACCCAGGTTACATCTACGGTCTCTTGGACCAAGATGTATCAATGATAACGA	663
Db	TGTATCAGTCCAGGATACATGTGTGACTGGCCCTGACCCGCAAAATTTCTGACCAACAACCT	256
OY	GAATGTATCAACAGATGATCTCTGTATCCCAACAAGAAATGCCAGCAAAAGNA	723
Db	GATCTCAAGAGAAATGACACTCCCTCAATCCACAAGGCAAAATGGGTACACTTGAAAC	196
OY	TACATTGTGTCTGTTTGTACTTGCTTTCTGTGATCTGCTCTTCAATACATACTGATGCC	783
Db	CTCATGGGCCCAAGTACGCTTCTTATGATGA---TGCGACCAAAATATGTATCTGGGGCA	139
OY	AGCTTACTGTTGATGTGGTTTCACTT	811
Db	GATTTAGGGTGAATGGCGGATACCTT	111

RESULT 10
CF824543
LOCUS
DEFINITION
CF824543 676 bp mRNA linear EST 01-APR-2000
EST701925 Coccidioides posadasii saprobic phase cDNA library, 2 to
4 Kb Coccidioides posadasii cDNA clone CIDAR90 5' end, mRNA
sequence.
CF824543
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
CF824543.1 GI:45930600
EST.
Coccidioides posadasii
Coccidioides posadasii
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Onygenales; mitosporic Onygenales; Coccidioides.
1 (bases 1 to 676)
Gardner,M.J. and Cole,G.T.
Analysis of gene expression in Coccidioides posadasii mycelia and
spherules via expressed sequence tags
Unpublished (2003)
Other_ESTs: EST701924

FEATURES	SOURCE	LOCATION/QUALIFIERS
1.	.676	
		/organism="Coccidioides posadaei"
		/mol_type="mRNA"
		/strain="C73"
		/db_xref="taxon:199306"
		/clone="CIDAR30"
		/dev_stage="saprobic phase (mycelia)"
		/lab_host="E. coli DH10B, T1 phage resistant"
		/clone_1fb="Coccidioides posadaei saprobic phase CDNA library 2 to 4 kb"
		/note="Vector: pExpress 1; Site 1: Not I; Site 2: Eco RV; Coccidioides posadaei saprobic phase CDNA library, size fractionated CDNA 2 to 4 kb"

ORIGIN	Query Match	13.0%	Score 106.4	DB 7	Length 676
Best Local Similarity	58.6%	Pred. No. 3.6e-19			
Matches 204	Conservative 0	Mismatches 141	Indels 3	Gaps 1	
Oy	464 TTGTTTGATTGGTTCTATGTCGTGGTCGCAATTCGACAGCAATCTCAAAACCAAGTTGCT				523
Db	121 TTGATTCATTGGAAGCAGATGTCGGTGCCATTGTAAAGTGCGCTCAGCGACAAAGGCGGT				180
Oy	524 ACAACATGTCGAAGCGGTGTTATCCATTGGGCTAAAGACTTTGGCTTGTGATGGGCTA				583

Db	181	ACAAATGCTGCAAAAGCGCTCGAATAAGGCACTGGCGTCTTCCCTGGGCAATTGAATGGGCA	240
Oy	584	AGTACAAATCTGAGTTAAATCTTAAACCGAGTTAATCTACGGCTCTTGAACAA	643
Db	241	GTCGGGAAATCAGAGTGAACGTATCACTCAGTCACGAGATACATGTGAAGCTGCCCTGACCCGA	300
Oy	644	ATGTTATCAATGTGTAACGAGAATTTGTAACAACAGATGATCTTGGTATCCACAACAA	703
Db	301	AGATTTCTTGAGACAAACCCCTATCTCCAAGGAAAGATGACCCCTCAATTCACAAGGCA	360
Oy	704	GAATGTCGGAACCAAGAAATATCAATGTGGCTGTTTCTACTCTTCTTGATCTGCTG	763
Db	361	AAATGGGTACCTGTGAAGACCTCAATGGGCCGAGTACGTTCTTATTAAGTGA---TGGA	417
Oy	764	CTTATACACTACTGCTGCGACAGCTTACTGCTGTGAATGAGGCTTCACTT	821
Db	418	GCAAAATATGTTACTGGGGCAGATTTGAAGGTTGAAGGGGCTACCTT	465

RESULT **11**

AJ638396 LOCUS 718 bp mRNA EST-05-MAY-2000
DEFINITION AJ638396 Mgc Mycosphaerella graminicola cDNA clone mgc06e02f, mRNA sequence.

ACCESSION AJ638396

VERSION AJ638396

KEYWORDS AJ638396.1 GI:47031453

SOURCE EST.
Mycosphaerella graminicola
Mycosphaerella graminicola
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes et Chaetothyriales; incertae sedis; Mycosphaerellaceae;
Mycosphaerella.
1 (bases 1 to 718)

REFERENCE Keon,J.P.R., Hagreaves,J.A., Antoniw,J.F. and Hammond-Kosack,K.
AUTHORS Analysis of expressed sequence tags from the wheat fungal leaf blotch pathogen, Mycosphaerella graminicola (anamorph Septoria tritici)
TITLE Fungal Genet. Biol. (2004) In press

JOURNAL COMMENT Contact: Keon J
Plant Pathogen Interactions Division,
Rothamsted Research,
Harpenden, Herts, UNITED KINGDOM
Tel.: +44(0)1582 763133
Fax: +44(0)1582 760981
Email: john.keon@bbc.ac.uk
Insert Length: 800 Std Error: 100.00
Seq primer: M13 reverse.

FEATURES	SOURCE
Location/Qualifiers	1. .718
	/organism="Mycosphaerella graminicola"
	/mol_type="mRNA"
	/strain="Scrlt"
	/db_xref="taxon:54734"
	/clone="mgc06602f"
	/clone_1fb="Mgc"
	/note="Vector: pSPORN1; library constructed from senseless wheat leaves 21-25 days after infection with Mycosphaerella graminicola exhibiting abundant hyphal growth and asexual sporulation"

Query Match	12.6%	Score 102.8;	DB 1;	Length 718;
Best Local Similarity	53.0%	Pred. No. 4e-18;		
Matches 268;	Conservative 0;	Mismatches 232;	Indels 6;	Gaps 2
QY 304	CAC TTG GTT AAC A CAG CGT GGT A C TGT G AAA A C T T C C A T G TGA A G T T A C C A G C C A G	363		
DB 49	C A C C T G C T C A C T T C C G C G C T T A C C G A G A A C T T G A C G C C A T G A G T A C C A C A C	108		
QY 364	A A G C T G A G A A G A T G T G A A G G T T A A C T G T T G G G T T C T T G T A T G T T C T C A G C C T T T	423		
DB 109	C G A T G C A G A A G C T T T G G G A G A T A A G T C G A C G C A C A T A T C A T C T G G T C T T T G C A C C G T G	168		

TITLE	A CDNA library prepared from <i>Fusarium graminearum</i> grown on a simple substrate
JOURNAL	Unpublished (2003)
COMMENT	Contact: Watson, Robert.J.

```

/120_just2: coli_dna5
/clone_11b="Pg09_AAPC_ECONC_Fusarium_graminearum_simple_substrata"
/notes="Vector: pBluescript II+, Site 1: EcoRI; Site 2: XhoI; Fusarium graminearum grown on a simple substrate-- minimal media supplemented with amino acids."

```

ORIGIN

Query Match	11.1%;	Score 90.8;	DB 7;	Length 892;
Best Local Similarity	54.5%;	Pred. No. 1.2e-14;		
Matches 204; Conservative	0;	Mismatches 167;	Indels 3;	Gaps 1;

QY	437	TCAAAGAAAGTATCAAGGTCCTTGTTGTTTTCATTGGTTCTATGTCGTGCATTG	496
DB	388	TCATGAGAGGTGCAAGGTGCCCGGTAGATTATGGTATTATGGAACATGTCGTGGTCATTG	329
QY	497	TCAAAGATCCCTCAAAAACCAAGTTGTCTCAACATGTCCAAAGGCTGATGTAATCCATTGG	556
DB	328	TCAAATGCCACAGCAACAAGCCCATATTAAGCGCGCAAAAGCAAGTGTTCACACCTGG	269
QY	557	CTAAGACTTTGGCTTTGATGATGGGCTAAGTACAAATCAGAGTTATTTCTTMAAACCAG	616
DB	268	CTGCTTCTCTTGCAGTGGAGTGGGCTCACGCTGGAAATCGGTGTCAACTGCATCTCTCTG	209
QY	617	GTTACATCTACGGTCCCTTTGACCAAGATGTTATCATGTGTATCGAAGATTGTACACA	676
DB	208	GCTATATGTGTGACTGTTTGACACAGAAATCTCGACCAACCTGTATTTAGAGAGA	149
QY	677	GATGATCTCTGATATCCCAACAAGAATGCCAACAAGGAATACATTGGGCTG	736
DB	148	CCCGACATCCCTCATCTCTCAGGGTCGATGGAGACTGCTTCAAGATTGAATGGAAACCG	89
QY	737	TTTTTGACTTGGCTTTCGATCTGCTGCTTCAATACACTATGGTGCACCTTATGGTTG	796
DB	88	TAACCTTTGTGATACGA---TGGCTATCTTATATGACTGGGCGAGATGTTGAGTTG	32
QY	797	ATGTCGTCTTACT	810
DB	31	ATGAGAGATACACT	18

Search completed: July 12, 2005, 14:20:51
Job time : 3392 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: July 9, 2005, 10:15:36 ; Search time 162 Seconds

(without alignments)
649.375 Million cell updates/sec

Title: US-10-720-018-2

Perfect score: 1415

Sequence: 1 MTDIYIPFRDGHULTVGA.....SAASYTGASLLVDGFTSW 272

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: GeneSeq_16Dec04:*
2: GeneSeq_1980s:*
3: GeneSeq_1990s:*
4: GeneSeq_2000s:*
5: GeneSeq_2001s:*
6: GeneSeq_2002s:*
7: GeneSeq_2003s:*
8: GeneSeq_2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1415	100.0	272	8	ADQ94602
2	424.5	30.0	258	3	AAV56815
3	402.5	28.4	257	8	ADN20270
4	396	28.0	211	8	ADS24003
5	393	27.8	266	5	AAE28702
6	381	26.9	283	8	ADJ98100
7	380	26.9	283	8	ADJ98118
8	378	26.7	283	4	AAW64777
9	378	26.7	283	4	AAW63579
10	378	26.7	283	8	ADJ98066
11	378	26.7	283	8	ADJ98066
12	378	26.7	283	8	ADJ98066
13	378	26.7	283	8	ADJ98066
14	376	26.6	283	8	ADJ98064
15	376	26.6	283	8	ADJ98104
16	376	26.6	283	8	ADJ98062
17	376	26.6	283	8	ADJ98112
18	374	26.4	245	8	ADN20651
19	374	26.4	283	8	ADJ98102
20	373	26.4	283	8	ADJ98108
21	373	26.4	283	8	ADJ98114
22	373	26.4	283	8	ADJ98116
23	370	26.1	283	8	ADJ98110
24	367	25.9	251	8	ADS25925
25	367	25.9	283	8	ADJ98106

26	364	25.7	263	6	ABJ26225
27	362	25.6	251	8	ADS25181
28	358.5	25.3	266	7	ADG25179
29	358.5	25.3	281	8	ADG25185
30	358.5	25.3	281	8	ADG25185
31	358.5	25.3	319	8	ADG24222
32	357	25.2	245	8	ADS25559
33	354.5	25.1	266	7	ADG25187
34	350	24.7	293	8	ADN21234
35	344	24.3	316	7	ABO66473
36	343.5	24.3	294	6	ABJ25625
37	343	24.2	268	8	ADS23247
38	340.5	24.1	257	5	ABP38717
39	340.5	24.1	257	8	ADG05249
40	337.5	23.9	267	7	ADG25186
41	337.5	23.9	267	7	ADG25184
42	336	23.7	255	8	ADN20291
43	326.5	23.1	284	8	ADN21004
44	324.5	22.9	245	8	ADN25229
45	322.5	22.8	256	7	ABO81728

ALIGNMENTS

RESULT 1	ADQ94602	standard; protein; 272 AA.
XX	ADQ94602	
AC	ADQ94602;	
DT	23-SEP-2004	(first entry)
DE	Ambrosiozyma monospora	NADH dependent L-xylose reductase.
KW	NADH dependent L-xylose reductase; fermentation; biomass; enzyme;	
KM	EC 1.1.1.10.	
OS	Ambrosiozyma monospora.	
XX	US2004132074-A1.	
PD	08-JUL-2004.	
PF	24-NOV-2003; 2003US-00720018.	
PR	16-FEB-2001; 2001FI-00000308.	
PR	15-FEB-2002; 2002WO-FI000125.	
PR	10-MAR-2003; 2003US-00257821.	
PR	12-SEP-2003; 2003FI-00001307.	
PA	(VALM) VALTTON TEKNIILINEN TUTKIMUSKESKUS.	
XX	Verho R, Richard P, Penttila M;	
PI	WPI; 2004-517001/49.	
DR	N-PSDB; ADQ94601.	
XX	New DNA molecule encoding NADH dependent L-xylose reductase, useful in	
PT	producing fermentation products or in conversion of cheap biomass to	
PT	useful products.	
PS	Claim 26; SEQ ID NO 2; 15pp; English.	
CC	The present invention relates to a NADH dependent L-xylose reductase	
CC	enzyme protein and its encoding polynucleotide. The invention is useful	
CC	in producing fermentation products or in conversion of cheap biomass to	
CC	useful products. The present sequence is an Ambrosiozyma monospora NADH	
CC	-1484 NADH dependent L-xylose reductase, EC 1.1.1.10.	
XX	Sequence 272 AA;	
SO	Query Match	100.0%; Score 1415; DB 8; Length 272;

Best Local Similarity 100.0%; Pred. No. 1,1e-137;
Matches 272; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTDTYIPFRPDGHLTIYTGACGGIAEALIKGLIAYGSDIALDIDOEKTAQKQAEYHKA 60
DB 1 MTDTYIPFRPDGHLTIYTGACGGIAEALIKGLIAYGSDIALDIDOEKTAQKQAEYHKA 60
QY 61 TEEIKLKEVPKMGSYACDISDPTVHKVFAQVADPGKLPPLHNTAGYCENPCEDYPA 120
DB 61 TEEIKLKEVPKMGSYACDISDPTVHKVFAQVADPGKLPPLHNTAGYCENPCEDYPA 120
QY 121 KNAEMKYNVLGSLIYVYSGAFAPKPLIKGASVVLIGSMGAIYVDPQNVYVYNSKA 180
DB 121 KNAEMKYNVLGSLIYVYSGAFAPKPLIKGASVVLIGSMGAIYVDPQNVYVYNSKA 180
QY 181 GVHIAKTLACEMAKNIRVNSLNPGYTGGPLTKRVINGNEELYNRWISGIPQORMSEPK 240
DB 181 GVHIAKTLACEMAKNIRVNSLNPGYTGGPLTKRVINGNEELYNRWISGIPQORMSEPK 240
QY 241 EYIGAVLYLSESAASYTTGASLVDGGFTSM 272
DB 241 EYIGAVLYLSESAASYTTGASLVDGGFTSM 272

RESULT 2

AA56815
ID AA56815 standard; protein; 258 AA.

XX AC AA56815;
XX AC
XX DT 31-MAR-2000 (first entry)
XX DE Bacillus D-arabinitol dehydrogenase.
XX KW Bacillus D-arabinitol dehydrogenase; clinical diagnosis; mycosis.
XX OS Bacillus sp.
XX PN JP11332569-A.
XX PD 07-DEC-1999.
XX PF 26-MAY-1998; 98JP-00143637.
XX PR 26-MAY-1998; 98JP-00143637.
XX PA (IKED-) IKEDA SHOKEN KK.
XX PA (NIPK) NIPPON KAYAKU KK.
XX DR WPI: 2000-091353/08.
XX DR N-PSDB; AA246762, AA246763.
XX PT Arabinitol dehydrogenase gene encoding D-arabinitol dehydrogenase -
XX PT useful as a clinical diagnosing agent for mycosis.
XX PS Claim 1; Page 10-11; 14pp; Japanese.
XX CC The invention relates to gene encoding D-arabinitol dehydrogenase,
CC isolated from Bacillus sp. IKD-5A868 strain. The protein can be expressed
CC by standard recombinant methodologies. D-arabinitol dehydrogenase is used
CC as a clinical diagnosing agent for mycosis. The present sequence
CC represents the D-arabinitol dehydrogenase enzyme
XX SQ Sequence 258 AA;

Query Match 30.0%; Score 424.5; DB 3; Length 258;
Best Local Similarity 38.0%; Pred. No. 5,1e-35;

Matches 103; Conservative 44; Mismatches 99; Indels 25; Gaps 7;

QY 8 PFDDGHLTIYTGACGGIAEALIKGLIAYGSDIALDI--DQEKTAQKQAEYHKAITEELK 65
DB 7 PFLDGVTVITVAGMGIGRAMMALALAQAGSHIVADIADRADEAQEA----- 52

QY 66 LKEVPKMGSYA---CDISDPTVHKVFAQVADPGKLPPLHNTAGYCENPCEDYPAK 121
DB 53 -NEIRSCGVAFATVQDVTDEAOVEALIRQVEOYGRLDV-LVNNAGIVHKDRAEELPYE 110
QY 122 MAEMKYNVLGSLIYVYSGAFAPKPLIKGASVVLIGSMGAIYVDPQNVYVYNSKAG 181
DB 111 TMRKVNWNVNSVFLMSKAGKVRIRQG-KG-STINISSGSLIVTPOQAAYVNSKAG 168
QY 182 VIHIAKTLACEMAKNIRVNSLNPGYTGGPLTKRVINGNEELYNRWISGIPQORMSEPK 241
DB 169 VIHIAKTLACEMAKNIRVNSLNPGYTGGPLTKRVINGNEELYNRWISGIPQORMSEPK 228
QY 242 YIGAVLYLSESAASYTTGASLVDGGFTSM 272
DB 229 LGIYAVLYASD-ASSFATGVFTIDGGYTTM 258

RESULT 3

ADN20270
ID ADN20270 standard; protein; 257 AA.

XX AC ADN20270;
XX AC
XX DT 02-DEC-2004 (first entry)
XX DE Bacterial polypeptide #2923.
XX OS

KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.

XX OS Bacteria.
XX PN US2003233675-A1.
XX PD 18-DEC-2003.
XX PF 20-FEB-2003; 2003US-00369493.
XX PR 21-FEB-2002; 2002US-0360039P.
XX PA (CAOY) CAO Y.
XX PA (HINK) HINKLE G J.
XX PA (SLAT) SLATER S C.
XX PA (CHEN) CHEN X.
XX PA (GOLD) GOLDMAN B S.
XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX DR WPI: 2004-061375/06.
XX PT New recombinant DNA construct comprising a promoter positioned to provide
XX PT for expression of a polynucleotide encoding a polypeptide from a
XX PT microbial source, useful for producing plants with improved properties.
XX PS Claim 1; SEQ ID NO 2923; 122pp; English.
XX CC The invention relates to a recombinant DNA construct comprising a
XX CC promoter functional in a plant cell, where the promoter is positioned to
XX CC provide for expression of a polynucleotide encoding a polypeptide from a
XX CC microbial source. The invention also relates to a transformed plant
XX CC comprising the recombinant DNA construct and a method of producing a
XX CC transformed plant having an improved property. The plant is a crop plant
XX CC such as maize or soybean. The method of producing a transformed plant
XX CC having an improved property comprises transforming a plant with the
XX CC recombinant DNA construct and growing the transformed plant, where the
XX CC polynucleotide or polypeptide is useful for improving plant properties.
XX CC The recombinant DNA construct is useful for producing plants with
XX CC improved plant properties, e.g. improved cold, heat or drought tolerance,

KM ethanol; enzyme.
 XX Hypocrea jecorina.
 OS
 XX WQ200266616-A2.
 PN
 XX 29-AUG-2002.
 PD
 XX
 PF 15-FEB-2002; 2002MO-F1000125.
 XX
 PR 16-FEB-2001; 2001FI-00000308.
 XX
 PA (VALM) VALTION TEKNIILLINEN TUTKIMUSKESKUS.
 PI Londeborough J, Penttilae M, Richard P;
 XX WPI; 2002-691618/74.
 DR N-PSDB; AAD46155.
 XX
 PT Genetically modified fungus for producing useful products such as
 PT ethanol, lactic acid and xylicol, from biomass containing L-arabinose,
 PT has increased ability to utilize L-arabinose.
 PS
 XX Disclosure; Fig 3; 32pp; English.
 CC The invention relates to genetically modified fungus with an increased
 CC ability to utilise L-arabinose, where the fungus has been transformed
 CC with a DNA sequence encoding an L-arabinol 4-dehydrogenase (EC 1.1.
 CC 1.12) or L-xylicolase reductase (EC 1.1.1.10) or both the DNA sequences.
 CC Genetically modified fungus is useful for producing useful products from
 CC biomass containing L-arabinose. The useful product include ethanol,
 CC lactic acid or xylicol preferably ethanol. It is also useful to ferment a
 CC carbon source such as biomass comprising agricultural or forestry
 CC products and waste products containing L-arabinose and also other
 CC pentoses or other fermentable sugars. The present sequence is T. reesei L
 CC -xylicolase reductase. (Updated on 29-AUG-2003 to standardise OS field)
 XX
 SQ Sequence 266 AA;
 Query Match 27.8%; Score 393; DB 5; Length 266;
 Best Local Similarity 36.4%; Pred. No. 9.9e-32;
 Matches 102; Conservative 51; Mismatches 103; Indels 24; Gaps 8;
 QY 1 MTDVPT-----FRPDGHLTVTGAGC--GLAEALIKGLIAYGSDIALDIDOEKTA 51
 DB 1 MPQVPTLNRLLDLFSLKGVVVTGASGPRGMEARCAEMKADAIITYSSKKEAE 60
 QY 52 KOAEYHKYATEBELKKEVPRKMSYACDISDPTVHKVPAQVAKDFGLPLHLVNTAGYCE 111
 DB 61 KNA-----EELTKYGVKVAVKVNQSDVNDVERFVNQVSDFGKIDAFIAN-AGATA 112
 QY 112 NPPCEDYPAKNAEKVKVKNLGLSYQAFAKPLIKEGKASVVLISMSGALYNDPON 171
 DB 113 NSGVVDGASDWDHVIQVDLSTGAYCAKAVGNAHFKQG--HGSIVITASMSGHVAANYQE 170
 QY 112 QVVVMSKAGVTHLAKTACCEWAKNIRVNSINPGYIVGPTKNTYNGINEBELYNMISGI 231
 DB 171 QTSIVVAKAGCTHILKRSLSLANEKRF-ARVNSISPGYIDGLSDFIDEKTEQEL---WRSMI 226
 QY 232 PQORMSEPEYIGAVLYLLESBAASYTTGASLLVDGFTS 271
 DB 227 PWGRNGDAKELKGAIVYLVS-ASSYTTGADIVIDGYTT 265
 RESULT 6
 ADJ98100 standard; protein; 283 AA.
 ADJ98100;
 06-MAY-2004 (first entry)
 Ketoreductase (KRED) protein SEQ ID NO:42.

XX 4-cyano-3-hydroxybutyric acid ester; 4-halo-3-hydroxybutyric acid ester;
 KM 4-halo-3-ketobutyric acid ester; halohydrin dehalogenase; cyanide;
 KM 4-nucleophile substituted 3-hydroxybutyric acid ester; amide;
 KM ketoreductase; cofactor; cofactor regeneration system; nucleophile;
 KM 4-nucleophile substituted-3-hydroxybutyric acid amide;
 KM 4-substituted 3-hydroxybutyric acid derivative; KRED; enzyme.
 OS
 XX Synthetic.
 PN WQ2004015132-A2.
 XX
 PD 19-FEB-2004.
 XX
 PF 11-AUG-2003; 2003MO-US025263.
 XX
 PR 09-AUG-2002; 2002US-00402436P.
 PR 11-AUG-2003; 2003US-00402436.
 XX
 PA (CODE-) CODEXIS INC.
 PI Davis SC, Grate JH, Gray DR, Gruber JM, Huismann GW, Ma SK;
 PI Newman LM, Sheldon R, Wang LA;
 XX WPI; 2004-238752/22.
 DR N-PSDB; ADJ98099.
 XX
 PT Producing 4-cyano-3-hydroxybutyric acid ester from 4-halo-3-
 PT hydroxybutyric acid ester, comprises contacting 4-halo-3-hydroxybutyric
 PT acid ester with halohydrin dehalogenase and cyanide.
 PS
 XX Example 9; SEQ ID NO 42; 168pp; English.
 CC The present invention describes a method (M1) for producing 4-cyano-3-
 CC hydroxybutyric acid ester (I) from a 4-halo-3-hydroxybutyric acid ester
 CC (II) or 4-halo-3-ketobutyric acid ester (III), comprising providing (II)
 CC or (III), and contacting (II) with a halohydrin dehalogenase and cyanide
 CC under conditions sufficient to form a reaction mixture for converting
 CC (II) to (I). Also described: (1) producing (M2) a 4-nucleophile
 CC substituted 3-hydroxybutyric acid ester or amide from (II) or amide,
 CC comprising: (a) providing (II), (III) or amide; and (b) contacting (II)
 CC or amide with a halohydrin dehalogenase and a nucleophile under
 CC conditions suitable to form a reaction mixture for converting (II) or
 CC amide to a 4-nucleophile substituted-3-hydroxybutyric acid or amide; or
 CC (c) contacting (III) or amide with a ketoreductase, a cofactor, a
 CC cofactor regeneration system, a nucleophile and a halohydrin dehalogenase
 CC to form a reaction mixture for converting (III) or amide to a 4-
 CC nucleophile substituted-3-hydroxybutyric acid ester or amide; and (2) a
 CC composition (IV) comprising halohydrin dehalogenase, a nucleophile, and
 CC (I) or amide. (M1) is useful for producing a 4-cyano-3-hydroxybutyric
 CC acid ester from a 4-halo-3-hydroxybutyric acid ester or 4-halo-3-
 CC ketobutyric acid ester. (M2) is useful for producing a 4-nucleophile
 CC substituted-3-hydroxybutyric acid ester or amide. (M1) and (M2) are
 CC useful for producing 4-substituted 3-hydroxybutyric acid derivatives. The
 CC present sequence represents a ketoreductase (KRED), which is used in the
 CC exemplification of the present invention.
 SQ Sequence 283 AA;
 Query Match 26.9%; Score 381; DB 8; Length 283;
 Best Local Similarity 36.1%; Pred. No. 1.9e-30;
 Matches 97; Conservative 52; Mismatches 98; Indels 22; Gaps 8;
 QY 8 FRPDGHLTVTGAGCGLAELIKGLIAYGSDIALDIDOEKTA KOAEYHKYATEBELK 67
 DB 29 FKLNGKVASITGSSSGIGTALAEAPQVADVAIWMYNSQDAGKAEALAKXGV----- 82
 QY 68 EVPKMSYACDISDPTVHKVPAQVAKDFGLPLHLVNTAG--YCENPCCEDYPAKNAEK 125
 DB 83 ---KVKAYKANVSSSDAVAKQTIEQIKDFGHLDIVVAN-AGIPWTKGAYITDDDDKHDPQ 138
 QY 126 MKKVNLLSLYS---QAFAPKPLIKEGKASVVLISMSGALYNDPONQVVVMSKXAG 181

Db 139 VIDVDLKGVGVAKHAGHRYERFEKEGKGA-LIFTASVSGHIVNIPOFQATYNAKAG 197
 Qy 182 VIHLAKTLACEMAKYNIRVNSLNPQYIGPLTKRVINGNEELNYRMISGIPQORMSEPK 241
 Db 198 VRHFAKSLAVEFAFP-ARVNSVSPGYINTEISDFV---PQRTQKMWSLVPLGRGETAE 253
 Qy 242 YIGAVLYLSSASAYTTGASLVDGGFT 270
 Db 254 LVGAYLFLASD-AGSYATGTDTIIVDGGYT 281

RESULT 7

ADJ98118 ADJ98118 standard; protein; 283 AA.

ADJ98118; ADJ98118;

06-MAY-2004 (first entry)

Ketoreductase (KRED) protein SEQ ID NO:60.

4-cyano-3-hydroxybutyric acid ester; 4-halo-3-hydroxybutyric acid ester;
 4-halo-3-ketobutyric acid ester; halohydrin dehalogenase; cyanide;
 4-nucleophile substituted 3-hydroxybutyric acid ester; amide;
 ketoreductase; cofactor; cofactor regeneration system; nucleophile;
 4-nucleophile substituted-3-hydroxybutyric acid amide;
 4-substituted 3-hydroxybutyric acid derivative; KRED; enzyme.

Synthetic.

WO2004015132-A2.

19-FEB-2004.

11-AUG-2003; 2003WO-US025263.

09-AUG-2002; 2002US-0402436P.

11-AUG-2003; 2003US-00402436.

(CODE-) CODEXIS INC.

David SC, Grate JH, Gray DR, Gruber JM, Huisman GW, Ma SK;
 Newman LM, Sheldon R, Wang LA;

WPI; 2004-238752/22.

N-PSDB; ADJ98117.

Producing 4-cyano-3-hydroxybutyric acid ester from 4-halo-3-hydroxybutyric acid ester, comprises contacting 4-halo-3-hydroxybutyric acid ester with halohydrin dehalogenase and cyanide.

Disclosure; SEQ ID NO 60; 168bp; English.

The present invention describes a method (M1) for producing 4-cyano-3-hydroxybutyric acid ester (I) from a 4-halo-3-hydroxybutyric acid ester (II) or 4-halo-3-ketobutyric acid ester (III), comprising providing (II) or (III), and contacting (II) with a halohydrin dehalogenase and cyanide under conditions sufficient to form a reaction mixture for converting (II) to (I). Also described: (1) producing (M2) a 4-nucleophile substituted 3-hydroxybutyric acid ester or amide from (II) or amide, comprising: (a) providing (II), (III) or amide; and (b) contacting (II) or amide with a halohydrin dehalogenase and a nucleophile under conditions suitable to form a reaction mixture for converting (II) or amide to a 4-nucleophile substituted-3-hydroxybutyric acid or amide; or (c) contacting (III) or amide with a ketoreductase, a cofactor, a cofactor regeneration system, a nucleophile and a halohydrin dehalogenase to form a reaction mixture for converting (III) or amide to a 4-nucleophile substituted-3-hydroxybutyric acid ester or amide; and (2) a composition (IV) comprising halohydrin dehalogenase, a nucleophile, and (I) or amide. (M1) is useful for producing a 4-cyano-3-hydroxybutyric acid ester from a 4-halo-3-hydroxybutyric acid ester or 4-halo-3-ketobutyric acid ester. (M2) is useful for producing a 4-nucleophile substituted-3-hydroxybutyric acid ester or amide. (M1) and (M2) are

CC useful for producing 4-substituted 3-hydroxybutyric acid derivatives. The present sequence represents a ketoreductase (KRED), which is used in the exemplification of the present invention.

Sequence 283 AA;

Query Match 26.9%; Score 381; DB 8; Length 283;
 Best Local Similarity 36.1%; Pred. No. 1.9e-30;
 Matches 97; Conservative 52; Mismatches 98; Indels 22; Gaps 8;

Qy 8 FRPDGHLITVYAGCGALAEALIKGLAVGSDIALDIDQETAKQKAEHYKATIEELTK 67
 Db 29 FLNKGKVASITGSSSGIGYALAEAPQGVADVATWNSQDRTGKAEALAKTYG----- 82
 Qy 68 EVPMKGSYACDISDPTVHKVFAQVAFGLPLHLVNTAG--YCENPCEBEPKAKNAEK 125
 Db 83 ---KTKAVKANVSSSDAKVKQITIEQIDFGHLDIVVAN-AGIPTKAGVITQDDDKHDDQ 138
 Qy 126 MKRVNLLGSLYVS---QAFKPLIKEGIGKASVVLIGSMGAIVNDPQNVVYNSKAG 181
 Db 139 VIDVDLKGVGVAKHAGHRYERFEKEGKGA-LIFTASVSGHIVNIPOFQATYNAKAG 197
 Qy 182 VIHLAKTLACEMAKYNIRVNSLNPQYIGPLTKRVINGNEELNYRMISGIPQORMSEPK 241
 Db 198 VRHFAKSLAVEFAFP-ARVNSVSPGYINTEISDFV---PQRTQKMWSLVPLGRGETAE 253
 Qy 242 YIGAVLYLSSASAYTTGASLVDGGFT 270
 Db 254 LVGAYLFLASD-AGSYATGTDTIIVDGGYT 281

RESULT 8

ADJ98098 ADJ98098 standard; protein; 283 AA.

ADJ98098; ADJ98098;

06-MAY-2004 (first entry)

Ketoreductase (KRED) protein SEQ ID NO:40.

4-cyano-3-hydroxybutyric acid ester; 4-halo-3-hydroxybutyric acid ester;
 4-halo-3-ketobutyric acid ester; halohydrin dehalogenase; cyanide;
 4-nucleophile substituted 3-hydroxybutyric acid ester; amide;
 ketoreductase; cofactor; cofactor regeneration system; nucleophile;
 4-nucleophile substituted-3-hydroxybutyric acid amide;
 4-substituted 3-hydroxybutyric acid derivative; KRED; enzyme.

Synthetic.

WO2004015132-A2.

19-FEB-2004.

11-AUG-2003; 2003WO-US025263.

09-AUG-2002; 2002US-0402436P.

11-AUG-2003; 2003US-00402436.

(CODE-) CODEXIS INC.

David SC, Grate JH, Gray DR, Gruber JM, Huisman GW, Ma SK;
 Newman LM, Sheldon R, Wang LA;

WPI; 2004-238752/22.

N-PSDB; ADJ98097.

Producing 4-cyano-3-hydroxybutyric acid ester from 4-halo-3-hydroxybutyric acid ester, comprises contacting 4-halo-3-hydroxybutyric acid ester with halohydrin dehalogenase and cyanide.

Disclosure; SEQ ID NO 40; 168bp; English.

XX 31-JAN-2001; 2001WO-US003186.
PE
XX
PR 31-JAN-2000; 2000US-0049421.
PR 08-DEC-2000; 2000US-00734237.
XX
PA (BIOC-) BIOCATALYTICS INC.
XX
PI Rozell DJ, Bui P, Hua L;
XX
DR WPI; 2001-483235/52.
DR N-PSDB; AAH74597.
XX
PT Designing synthetic nucleic acid sequences for improved amplification,
PT expression in host cell, by comparing free energy of folding of a
PT starting polynucleotide and a modified polynucleotide having a codon
PT replacement.
XX
XX
PS Claim 6; Page 103; 117pp; English.
XX
XX The present sequence is a NADPH-dependent carbonyl reductase. The
CC polynucleotide was modified using the method of the invention. The
CC specification describes a method for designing a synthetic
CC polynucleotide. The method comprises providing a starting polynucleotide,
CC determining the predicted free energy of folding per base of the
CC polynucleotide, modifying the polynucleotide by replacing a codon with a
CC different codon to provide a modified polynucleotide, determining free
CC energy of folding per base of the modified polynucleotides, and comparing
CC this with that of the original polynucleotide. The method is useful for
CC developing nucleic acid sequences that enhance expression of the encoded
CC protein in a heterologous host. The design and preparation of the
CC synthetic genes are used in application of gene shuffling, directed
CC evolution and molecular breeding methods. The method allows expression of
CC genes from various organisms such as mammals, plants, yeast, fungi and
CC bacteria in prokaryotic hosts, such as *Escherichia coli* and eukaryotic
CC hosts at commercially viable levels, in particular proteins with low
CC yield such as methionine gamma-lyase from *P. putida*
XX
SQ Sequence 283 AA;
XX
Query Match 26.7%; Score 378; DB 4; Length 283;
Best Local Similarity 36.8%; Pred. No. 3.9e-30;
Matches 99; Conservative 48; Mismatches 100; Indels 22; Gaps 8;
QY 8 FRPDGHTITVAGCGLAELIKGLAVGSDIALLDIDOEKTAQAQEHYKATBELK 67
DB 29 FKUNGKVASITGSSSGIGYALAEAFQVGDVAIWNVSHDRTGAKALAKYGV----- 82
QY 68 EVPRMGSYACDISDPTVHVAFQVAKDFGKLPLHLVNTAG--YCENFPCEDPAPAKAEK 125
DB 83 ---KVKAAYKANVSSSDAVKQTIHQIYKDFGLHDIIVAN--AGIPWTKGAYIDQDDDKHFDQ 138
QY 126 MVKVNLLGSLYVS---QAFAPKPLIKEGIGKASVVLGSMGALVNDPQNVVYNNMSKAG 181
DB 139 VVDVDLGVGVAAAGHGFHFERERKEGKKA-LVFTASMSGHIVNVPQFATYNAKAG 197
QY 182 VIHLAKTACEMAKTNIRVSNLNGPYTYGPTKAVINGNEELVYRWISGIPQOQMSBPKE 241
DB 198 VRHFAKSLAVEFAF-ARVSVSBCYITNTEISDFV---PORTQKMSVLVLGSGGETAB 253
QY 242 YIGAVLYLSESAAVYTTGASLVDGGET 270
DB 254 LVGAFLFLASD-AGSVATGTDIIVDGCVT 281
XX
RESULT 11
ADJ98066
ID ADJ98066 standard; protein; 283 AA.
XX
XX ADJ98066;
XX
XX 06-MAY-2004 (first entry)

DE Ketoreductase 4 (KRED CR2-4) protein SEQ ID NO:8.
XX
XX 4-cyano-3-hydroxybutyric acid ester; 4-halo-3-hydroxybutyric acid ester;
KM 4-halo-3-ketobutyric acid ester; haloaldrin dehalogenase; cyanide;
KM 4-nucleophile substituted 3-hydroxybutyric acid ester; amide;
KM ketoreductase; cofactor; cofactor regeneration system; nucleophile;
KM 4-nucleophile substituted 3-hydroxybutyric acid amide;
KM 4-substituted 3-hydroxybutyric acid derivative; KRED; enzyme.
XX
OS Candida magnoliae.
XX
XX WO2004015132-A2.
XX
XX 19-FEB-2004.
XX
XX 11-AUG-2003; 2003WO-US025263.
XX
XX 09-AUG-2002; 2002US-0402436P.
XX 11-AUG-2003; 2003US-00402436.
XX
XX (CODE-) CODEXIS INC.
XX
XX Davis SC, Grate JH, Gray DR, Gruber JM, Huisman GW, Ma SK;
PI Newman LM, Sheldon R, Wang LA;
XX
XX WPI; 2004-238752/22.
DR N-PSDB; ADJ98065.
XX
XX Producing 4-cyano-3-hydroxybutyric acid ester from 4-halo-3-
PT hydroxybutyric acid ester, comprises contacting 4-halo-3-hydroxybutyric
PT acid ester with haloaldrin dehalogenase and cyanide.
XX
XX Example 1, SEQ ID NO 8; 16pp; English.
XX
XX The present invention describes a method (M1) for producing 4-cyano-3-
CC hydroxybutyric acid ester (I) from a 4-halo-3-hydroxybutyric acid ester
CC (II) or 4-halo-3-ketobutyric acid ester (III), comprising providing (II)
CC or (III), and contacting (II) with a haloaldrin dehalogenase and cyanide
CC under conditions sufficient to form a reaction mixture for converting
CC (II) to (I). Also described: (1) producing (M2) a 4-nucleophile
CC substituted 3-hydroxybutyric acid ester or amide from (II) or amide,
CC comprising: (a) providing (II), (III) or amide; and (b) contacting (II)
CC or amide with a haloaldrin dehalogenase and a nucleophile under
CC conditions suitable to form a reaction mixture for converting (II) or
CC amide to a 4-nucleophile substituted 3-hydroxybutyric acid or amide; or
CC (c) contacting (III) or amide with a ketoreductase, a cofactor, a
CC cofactor regeneration system, a nucleophile and a haloaldrin dehalogenase
CC to form a reaction mixture for converting (III) or amide to a 4-
CC nucleophile substituted 3-hydroxybutyric acid ester or amide; and (2) a
CC composition (IV) comprising haloaldrin dehalogenase, a nucleophile, and
CC (1) or amide. (M1) is useful for producing a 4-cyano-3-hydroxybutyric
CC acid ester from a 4-halo-3-hydroxybutyric acid ester or 4-halo-3-
CC ketobutyric acid ester. (M2) is useful for producing a 4-nucleophile
CC substituted 3-hydroxybutyric acid ester or amide. (M1) and (M2) are
CC useful for producing 4-substituted 3-hydroxybutyric acid derivatives. The
CC present sequence represents a ketoreductase (KRED), which is used in the
CC exemplification of the present invention.
XX
SQ Sequence 283 AA;
XX
Query Match 26.7%; Score 378; DB 8; Length 283;
Best Local Similarity 36.8%; Pred. No. 3.9e-30;
Matches 99; Conservative 48; Mismatches 100; Indels 22; Gaps 8;
QY 8 FRPDGHTITVAGCGLAELIKGLAVGSDIALLDIDOEKTAQAQEHYKATBELK 67
DB 29 FKUNGKVASITGSSSGIGYALAEAFQVGDVAIWNVSHDRTGAKALAKYGV----- 82
QY 68 EVPRMGSYACDISDPTVHVAFQVAKDFGKLPLHLVNTAG--YCENFPCEDPAPAKAEK 125
DB 83 ---KVKAAYKANVSSSDAVKQTIHQIYKDFGLHDIIVAN--AGIPWTKGAYIDQDDDKHFDQ 138
QY 126 MVKVNLLGSLYVS---QAFAPKPLIKEGIGKASVVLGSMGALVNDPQNVVYNNMSKAG 181

XX The present invention describes a method (M1) for producing 4-cyano-3-hydroxybutyric acid ester (I) from a 4-halo-3-hydroxybutyric acid ester (II) or 4-halo-3-ketobutyric acid ester (III), comprising providing (II) or (III), and contacting (II) with a haloaldehyde dehydrogenase and cyanide under conditions sufficient to form a reaction mixture for converting (II) to (I). Also described: (1) producing (M2) a 4-nucleophile substituted 3-hydroxybutyric acid ester or amide from (II) or amide, comprising: (a) providing (II), (III) or amide; and (b) contacting (II) or amide with a haloaldehyde dehydrogenase and a nucleophile under conditions suitable to form a reaction mixture for converting (II) or amide to a 4-nucleophile substituted 3-hydroxybutyric acid ester or amide; or (c) contacting (III) or amide with a ketoreductase, a cofactor, a cofactor regeneration system, a nucleophile and a haloaldehyde dehydrogenase to form a reaction mixture for converting (III) or amide to a 4-nucleophile substituted 3-hydroxybutyric acid ester or amide; and (2) a composition (IV) comprising haloaldehyde dehydrogenase, a nucleophile, and (I) or amide. (M1) is useful for producing a 4-cyano-3-hydroxybutyric acid ester from a 4-halo-3-hydroxybutyric acid ester or 4-halo-3-ketobutyric acid ester. (M2) is useful for producing a 4-nucleophile substituted 3-hydroxybutyric acid ester or amide. (M1) and (M2) are useful for producing 4-substituted 3-hydroxybutyric acid derivatives. The present sequence represents a ketoreductase (KRED), which is used in the exemplification of the present invention.

SQ Sequence 283 AA;

Query Match 26.7%; Score 378; DB 8; Length 283;
Best Local Similarity 36.8%; Pred. No. 3.9e-30;
Matches 99; Conservative 48; Mismatches 100; Indels 22; Gaps 8;

QY 8 FRPDGHLITVTCAGCGLAELIKGLAYGSDIALDIDQETPAKQAEYHKYATELKLK 67
DB 29 FDLNKGVALVTCAGCGIGYALAEAFQVQADVAIWNHSHDAGKAEALAKTKGV----- 82
QY 68 EYPMKGSYACDISDSDYTHKVPQAVKADGKPLHLVNTAG--YCNFPCEDYPAKNAEK 125
DB 83 ---KVKAKYANVSSDAVAKQTTEQOI KDFGHLIDIVAN--AGIPWKGAVIDDDHGFPOQ 138
QY 126 MYKNILGSLVYS----QAFAPKLIKEGIKGASVVLGSGMGAIVNDPONVYNNMSKAG 181
DB 139 VVDVVLKGVGVYAKKAGRHFRFRFEKGGKGA-LFTTASMSGHIVAVPOFATYNAKAG 197
QY 182 VIHLAKTLACEMAKYINRVNSLMPGTYGPTLTKVINGNEBELYNEMISGIPQORMSEPK 241
DB 198 VTHFPAKSLAVEPAP-ARVNSVSPGYINTEISDFV---PQETQNMWMSLVPLGRGETAB 253
QY 242 YIGANVYLSSASAYTTGASLLVDGGFT 270
DB 254 LVGAYLFLASD-AGSYATGTDTIIVDGGYT 281

RESULT 14

ADS22515 standard; protein; 253 AA.

ADS22515;

02-DEC-2004 (first entry)

Bacterial polypeptide #11548.

Recombinant DNA construct; transformed plant; improved plant property;
cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic;
pathogen tolerance; pest tolerance; plant disease resistance;
cell cycle pathway modification; plant growth regulator;
homologous recombination; seed oil yield; protein yield; carbohydrate;
nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
bacterial polypeptide.

Bacteria.

US2003233675-A1.

XX 18-DEC-2003.
PD 20-FEB-2003; 2003US-00369493.
XX 21-FEB-2002; 2002US-0360039P.
PR (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
DR WPI; 2004-061375/06.
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
PS Claim 1, SEQ ID NO 11548; 122pp; English.

The invention relates to a recombinant DNA construct comprising a promoter functional in a plant cell, where the promoter is positioned to provide for expression of a polynucleotide encoding a polypeptide from a microbial source. The invention also relates to a transformed plant comprising the recombinant DNA construct and a method of producing a transformed plant having an improved property. The plant is a crop plant such as maize or soybean. The method of producing a transformed plant having an improved property comprises transforming a plant with the recombinant DNA construct and growing the transformed plant, where the polynucleotide or polypeptide is useful for improving plant properties. The recombinant DNA construct is useful for producing plants with improved plant properties, e.g. improved cold, heat or drought tolerance, tolerance to herbicides, extreme osmotic conditions, pathogens or pests, increased resistance to plant disease, better growth rate by modification of the cell cycle pathway with plant growth regulators, increased rate of homologous recombination, modified seed oil or protein yield and/or content, improved yield by modification of carbohydrate, nitrogen or phosphorus use and/or uptake, by modification of photosynthesis or by providing improved plant growth and development under at least one stress condition, improved lignin production or improved galactomannan production. This sequence represents a bacterial polypeptide used in the scope of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html.

SQ Sequence 253 AA;

Query Match 26.6%; Score 376; DB 8; Length 253;
Best Local Similarity 35.8%; Pred. No. 5.3e-30;
Matches 96; Conservative 43; Mismatches 109; Indels 20; Gaps 7;

QY 8 FRPDGHLITVTCAGCGLAELIKGLAYGSDIALDIDQETPAKQAEYHKYATELKLK 67
DB 3 FDLNKGVALVTCAGCGIGYALAEAFQVQADVAIWNHSHDAGKAEALAKTKGV----- 57
QY 68 EYPMKGSYACDISDSDYTHKVPQAVKADGKPLHLVNTAGYCNFPCEDYPAKNAEK 127
DB 58 ---KSIQTAADVTSKQALTDVAIVARTQAEIAGLSL--AVNAAGIANNPAEMSEESQFTWM 113
QY 128 MYKNILGSLVYSQAAPKPLIKEGIKGASVVLGSGMGAIVNDPONVYNNMSKAGYTHLAK 187
DB 114 DINLKGVLSCQAEENNAWLKNG-KGA-IVNTASMSGVTVNGLMCHYNASKAGVTHMSK 171
QY 188 TIACEMAKYINRVNSLMPGTYGPTLTKVINGNEBELYNR---WISGIPQORMSEPK 244
DB 172 SMAEMWVGRIKRVITISPGYATDM-----WIREMWHQITLFBQETQMGMAGADVAVG 226
QY 245 AVLYLSSASAYTTGASLLVDGGFTSW 272
DB 227 PAIFLLSD-AASFVTVGVDLVDGGFCFW 253

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 10:21:36 ; Search time 43 Seconds

(without alignments)
472.199 Million cell updates/sec

Title: US-10-720-018-2

Perfect score: 1415

Sequence: 1 MTDYPTFRFGHLLTVTGA.....SAASYTTGASLLVDGFTSW 272

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	686	48.5	231	US-08-336-198C-7	Sequence 7, Appli
2	686	48.5	231	US-09-184-965-7	Sequence 7, Appli
3	586	41.4	282	US-09-248-796A-14132	Sequence 14132, A
4	378	26.7	283	US-09-367-012-1	Sequence 1, Appli
5	378	26.7	283	US-09-777-157A-1	Sequence 70, Appli
6	378	26.7	283	US-09-734-237B-70	Sequence 16593, A
7	372.5	26.3	308	US-09-248-796A-16593	Sequence 12990, A
8	344	24.3	316	US-09-489-039A-12990	Sequence 3542, Ap
9	340.5	24.1	257	US-09-134-001C-3562	Sequence 30474, A
10	340.5	22.8	256	US-09-252-991A-30474	Sequence 6, Appli
11	321.5	22.7	262	US-09-363-189B-6	Sequence 2, Appli
12	312.5	22.1	292	US-09-468-738A-2	Sequence 2, Appli
13	312.5	22.1	292	US-09-940-019-2	Sequence 2, Appli
14	312.5	22.1	292	US-09-940-037A-2	Sequence 2, Appli
15	312.5	22.1	296	US-09-468-738A-23	Sequence 23, Appli
16	312.5	22.1	296	US-09-940-019-23	Sequence 23, Appli
17	312.5	22.1	296	US-09-940-037A-23	Sequence 16, Appli
18	307	21.7	251	US-09-922-501-16	Sequence 14805, Ap
19	302	21.3	275	US-09-107-532A-4805	Sequence 14861, A
20	296	20.9	249	US-09-902-540-14861	Sequence 12, Appli
21	290	20.5	258	US-09-504-358-12	Sequence 12, Appli
22	290	20.5	258	US-09-954-314-12	Sequence 7, Appli
23	290	20.5	258	US-10-230-562-12	Sequence 7, Appli
24	289	20.4	256	US-08-594-808B-7	Sequence 16592, A
25	282	19.9	290	US-09-248-796A-16592	Sequence 1, Appli
26	279	19.7	303	US-09-002-298-1	Sequence 1, Appli
27	279	19.7	303	US-09-481-277-1	Sequence 1, Appli

28	279	19.7	322	US-09-634-955B-11	Sequence 11, Appli
29	279	19.7	322	US-09-816-760-11	Sequence 11, Appli
30	279	19.7	322	US-09-838-561-11	Sequence 11, Appli
31	273.5	19.3	290	US-09-248-796A-16757	Sequence 16757, A
32	270.5	19.1	269	US-09-134-000C-4929	Sequence 4929, Ap
33	268	18.9	262	US-09-328-352-5742	Sequence 6693, Ap
34	265.5	18.8	253	US-09-543-681A-6693	Sequence 41628, A
35	265.5	18.8	276	US-09-270-767-41628	Sequence 4431, Ap
36	264.5	18.7	274	US-09-134-001C-4431	Sequence 382, Ap
37	264.5	18.7	306	US-09-710-279-382	Sequence 7731, Ap
38	263	18.6	261	US-09-328-352-7731	Sequence 25, Appli
39	262	18.5	303	US-09-816-760-13	Sequence 13, Appli
40	262	18.5	303	US-09-838-561-13	Sequence 13, Appli
41	262	18.5	303	US-09-107-433-3795	Sequence 3795, Ap
42	261	18.4	254	US-09-978-758-2	Sequence 2, Appli
43	260.5	18.4	254	US-09-489-039A-11936	Sequence 11936, A
44	259.5	18.3	259	US-09-489-039A-7802	Sequence 7802, Ap
45	257.5	18.2	308	US-09-489-039A-7802	Sequence 7802, Ap

ALIGNMENTS

RESULT 1
US-08-336-198C-7
; Sequence 7, Application US/08336198C
; Patent No. 5866382
; GENERAL INFORMATION:
; APPLICANT: Hallborn, Johan
; APPLICANT: Penttila, Merja
; APPLICANT: Ojamo, Heikki
; APPLICANT: Keranen, Sirkka
; APPLICANT: Hahn-Hagerdal, Barbel
; APPLICANT: Waldfriedson, Mats
; APPLICANT: Aliraksinen, Ulla
; TITLE OF INVENTION: XYLOSE UTILIZATION BY RECOMBINANT YEASTS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 301 N. Washington St.
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22046-3487
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336.198C
; FILING DATE: 03-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 231 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-336-198C-7
Query Match 48.5% Score 686; DB 2; Length 231;
Best Local Similarity 55.6%; Pred. No. 2.2e-67;
Matches 130; Conservative 49; Mismatches 49; Indels 6; Gaps 3;
QY 42 LDIDOEKTAQAQAEYHKYATEBLTKVEYKMG---SYACDISDSDDTVAKFAQVAKDQFGK 98

Db 1 MNLERTSAK--EVLGGEETLKGHNSAIGQVSAMSCNIGDAEADATFSSINEHKG 58
Qy 99 LPLHLVNTAGYCNFPCEDYPAKNAEKVKNLGLSVSOAFKPLIKEIKKASVLI 158
Db 59 IADLLINTAGYCNFPATYPAATNAESIMKVNGLGSFYVSOFPARPLONNLRG-STILI 117
Qy 159 GMSGAIYVNDPONOVVYMSKAGVIHLAKTLACEMAKYINRVNSLNPGYIYGPLTKVYN 218
Db 118 GMSGAIYVNDPONOVVYMSKAGVIHLAKTLACEMAKYINRVNSLNPGYIYGPLTKVYN 177
Qy 219 GNEELYNRMISGIPQORMSEPKYIGAVLYLSESASVTTGASLVDGFTSW 272
Db 178 GHEMKEMESKIPMKRMAEPKEFVGSILYLASETASVTTGHNLVVDGGYECW 231

RESULT 2
US-09-184-965-7
Sequence 7, Application US/09184965
Patent No. 6582944
GENERAL INFORMATION:
APPLICANT: Hallborn, Johan
APPLICANT: Penttila, Merja
APPLICANT: Ojamo, Heikki
APPLICANT: Keranen, Sirkka
APPLICANT: Hahn-Hagerdal, Barbel
APPLICANT: Walderidsson, Mats
APPLICANT: Alakainen, Ulla
TITLE OF INVENTION: Xylose UTILIZATION BY RECOMBINANT YEASTS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: 301 N. Washington St.
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22046-3487
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/184,965
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/336,198
FILING DATE: 03-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M.
REGISTRATION NUMBER: 28,977
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-205-8000
TELEFAX: 703-205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 231 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-184-965-7

Query Match 48.5%; Score 686; DB 4; Length 231;
Best Local Similarity 55.6%; Pred. No. 2.2e-67;
Matches 130; Conservative 49; Mismatches 49; Indels 6; Gaps 3;

Qy 42 LIDDEKTAQAQAEYHYKATVEKJKEVDMKG---SYACDISDPTVHKVPQVAKDFK 98
Db 1 MNLERTSAK--EVLGGEETLKGHNSAIGQVSAMSCNIGDAEADATFSSINEHKG 58
Qy 99 LPLHLVNTAGYCNFPCEDYPAKNAEKVKNLGLSVSOAFKPLIKEIKKASVLI 158

Db 59 IADLLINTAGYCNFPATYPAATNAESIMKVNGLGSFYVSOFPARPLONNLRG-STILI 117
Qy 159 GMSGAIYVNDPONOVVYMSKAGVIHLAKTLACEMAKYINRVNSLNPGYIYGPLTKVYN 218
Db 118 GMSGAIYVNDPONOVVYMSKAGVIHLAKTLACEMAKYINRVNSLNPGYIYGPLTKVYN 177
Qy 219 GNEELYNRMISGIPQORMSEPKYIGAVLYLSESASVTTGASLVDGFTSW 272
Db 178 GHEMKEMESKIPMKRMAEPKEFVGSILYLASETASVTTGHNLVVDGGYECW 231

RESULT 3
US-09-248-796A-14132
Sequence 14132, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstein et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
PRIOR FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 14132
LENGTH: 282
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-14132

Query Match 41.4%; Score 586; DB 4; Length 282;
Best Local Similarity 54.6%; Pred. No. 3.5e-56;
Matches 113; Conservative 42; Mismatches 48; Indels 4; Gaps 2;

Qy 5 IPRFRFDGHLITVYAGCGLAELIKGLLAGSDIALDIDQETAAKQAEYHYKATVEEL 64
Db 73 VPSFRIDGKLVITLGGSGDLAAVSRALLAKGADVALVDMNLERTQOAAADVLQMGEGM 132
Qy 65 KLR--EYPMKGSYACDISDSDTYHKVPQVAKDFGLPLHLVNTAGYCNFPCEDYPAK 121
Db 133 GKXESPIGQVSAMSCNIGDAEADLTFKALNEHKGKISSVLVNTAGYAEVPAEEYPAK 192
Qy 122 NAEKVVKNLGLSVSOAFKPLIKEIKKASVLIIGMSGAIYVNDPONOVVYMSKAG 181
Db 193 NAEKVVKNLGLSVSOAFKPLIKEIKKASVLIIGMSGAIYVNDPONOVVYMSKAG 251
Qy 182 VIHLAKTLACEMAKYINRVNSLNPGYI 208
Db 252 VIHLAKTLACEMAKYINRVNSLNPGYI 278

RESULT 4
US-09-367-012-1
Sequence 1, Application US/09367012
Patent No. 6218156
GENERAL INFORMATION:
APPLICANT: Yasohara, Yoshihiko
APPLICANT: Kizaki, No. 6218156Iyuki
APPLICANT: Hasegawa, Junzo
APPLICANT: Wada, Masaru
APPLICANT: Shimizu, Sakayu
APPLICANT: Kataoka, Michihiko
APPLICANT: Yamamoto, Kazuhiko
APPLICANT: Kawabata, Hiroshi
APPLICANT: Kita, Keiko
TITLE OF INVENTION: No. 6218156el carbonyl reductase, gene coding same, and method
FILE REFERENCE: S72030PCUS
CURRENT APPLICATION NUMBER: US/09/367,012

CURRENT FILING DATE: 1999-11-24
EARLIER APPLICATION NUMBER: PCT/JP97/03051
EARLIER FILING DATE: 1997-09-01
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 1
LENGTH: 283
TYPE: PRT
ORGANISM: Candida magnoliae
US-09-367-012-1

Query Match 26.7%; Score 378; DB 3; Length 283;
Best Local Similarity 36.8%; Pred. No. 3.7e-33;
Matches 99; Conservative 48; Mismatches 100; Indels 22; Gaps 8;

QY 8 FRPDGHLITVITGACGGLAELIKGLAYGSDIALDIDOEKTAQKQAEYHKYATEELKTK 67
DB 29 FKLNGKVASITGSSGIGALAEAFQVADVAIWNSHDAGKALAKKYG----- 82
QY 68 EVPKKGSTACDISDSDYHKKVFAQVAKDFGLPLHLVNTAG--YCENPCECDYPAKNAEK 125
DB 83 ---KVKAYKANVSSSDAVKQTEEQIDFGHLDIVAN-AGIPWTKGAYIDQDDDKHPDQ 138
QY 126 MVKNVNLGSLVYS---QAFAPLKEGKGSVVLIGMSGATVNDPQNVVYNNMSKAG 181
DB 139 VVDVLDKGVYAKIAGHFRERFEKEGKGA-LVFTASMSGHIVNVQFOATYNAAKAG 197
QY 182 VIHLAKTACEMAKYNIRVNSLNPYIYGPLTKVINGNEELYNRMISGIPQORMSEPK 241
DB 198 VRHFAKSLAVEFAPF-ARVNSVSPGYINTBISDFV---PQETQKNMWSLVPLGRGETAE 253
QY 242 YIGAVLYLSSASAYTTGASLVDGFT 270
DB 254 LVGAYLFLASD-AGSYATGTDIIVDGGYT 281

RESULT 5

US-09-777-157A-1
Sequence 1, Application US/09777157A
Patent No. 6448052
GENERAL INFORMATION:
APPLICANT: Yasohara, Yoshihiko
APPLICANT: Kizaki, No. 6448052yuki
APPLICANT: Hasegawa, Junzo
APPLICANT: Wada, Masaru
APPLICANT: Shimizu, Sakayu
APPLICANT: Katoko, Michihiko
APPLICANT: Yamamoto, Kazuhiko
APPLICANT: Kawabata, Hiroshi
APPLICANT: Kita, Keiko
TITLE OF INVENTION: Carbonyl Reductase Enzyme and Methods for its Use
FILE REFERENCE: 068383.0110
CURRENT APPLICATION NUMBER: US/09/777,157A
CURRENT FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: US 09/367,012
PRIOR FILING DATE: 1999-11-24
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patent In version 3.0
SEQ ID NO 1
LENGTH: 283
TYPE: PRT
ORGANISM: Candida magnoliae
US-09-777-157A-1

Query Match 26.7%; Score 378; DB 4; Length 283;
Best Local Similarity 36.8%; Pred. No. 3.7e-33;
Matches 99; Conservative 48; Mismatches 100; Indels 22; Gaps 8;

QY 8 FRPDGHLITVITGACGGLAELIKGLAYGSDIALDIDOEKTAQKQAEYHKYATEELKTK 67
DB 29 FKLNGKVASITGSSGIGALAEAFQVADVAIWNSHDAGKALAKKYG----- 82
QY 68 EVPKKGSTACDISDSDYHKKVFAQVAKDFGLPLHLVNTAG--YCENPCECDYPAKNAEK 125

DB 83 ---KVKAYKANVSSSDAVKQTEEQIDFGHLDIVAN-AGIPWTKGAYIDQDDDKHPDQ 138
QY 126 MVKNVNLGSLVYS---QAFAPLKEGKGSVVLIGMSGATVNDPQNVVYNNMSKAG 181
DB 139 VVDVLDKGVYAKIAGHFRERFEKEGKGA-LVFTASMSGHIVNVQFOATYNAAKAG 197
QY 182 VIHLAKTACEMAKYNIRVNSLNPYIYGPLTKVINGNEELYNRMISGIPQORMSEPK 241
DB 198 VRHFAKSLAVEFAPF-ARVNSVSPGYINTBISDFV---PQETQKNMWSLVPLGRGETAE 253
QY 242 YIGAVLYLSSASAYTTGASLVDGFT 270
DB 254 LVGAYLFLASD-AGSYATGTDIIVDGGYT 281

RESULT 6

US-09-734-237B-70
Sequence 70, Application US/09734237B
Patent No. 6818752
GENERAL INFORMATION:
APPLICANT: Rozzelli, J. David
APPLICANT: But, Peter
APPLICANT: Hua, Ling
TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
FILE REFERENCE: B583:40608
CURRENT APPLICATION NUMBER: US/09/734,237B
CURRENT FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 09/494,921
PRIOR FILING DATE: 2000-01-31
NUMBER OF SEQ ID NOS: 79
SOFTWARE: Patent In version 3.1
SEQ ID NO 70
LENGTH: 283
TYPE: PRT
ORGANISM: Candida magnoliae
US-09-734-237B-70

Query Match 26.7%; Score 378; DB 4; Length 283;
Best Local Similarity 36.8%; Pred. No. 3.7e-33;
Matches 99; Conservative 48; Mismatches 100; Indels 22; Gaps 8;

QY 8 FRPDGHLITVITGACGGLAELIKGLAYGSDIALDIDOEKTAQKQAEYHKYATEELKTK 67
DB 29 FKLNGKVASITGSSGIGALAEAFQVADVAIWNSHDAGKALAKKYG----- 82
QY 68 EVPKKGSTACDISDSDYHKKVFAQVAKDFGLPLHLVNTAG--YCENPCECDYPAKNAEK 125
DB 83 ---KVKAYKANVSSSDAVKQTEEQIDFGHLDIVAN-AGIPWTKGAYIDQDDDKHPDQ 138
QY 126 MVKNVNLGSLVYS---QAFAPLKEGKGSVVLIGMSGATVNDPQNVVYNNMSKAG 181
DB 139 VVDVLDKGVYAKIAGHFRERFEKEGKGA-LVFTASMSGHIVNVQFOATYNAAKAG 197
QY 182 VIHLAKTACEMAKYNIRVNSLNPYIYGPLTKVINGNEELYNRMISGIPQORMSEPK 241
DB 198 VRHFAKSLAVEFAPF-ARVNSVSPGYINTBISDFV---PQETQKNMWSLVPLGRGETAE 253
QY 242 YIGAVLYLSSASAYTTGASLVDGFT 270
DB 254 LVGAYLFLASD-AGSYATGTDIIVDGGYT 281

RESULT 7

US-09-248-796A-16593
Sequence 16593, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Ketch Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A

:
 : CURRENT FILING DATE: 1999-02-12
 : PRIOR APPLICATION NUMBER: US 60/074,725
 :
 : PRIOR FILING DATE: 1998-02-13
 : PRIOR APPLICATION NUMBER: US 60/096,409
 : PRIOR FILING DATE: 1998-08-13
 : NUMBER OF SEQ ID NOS: 28208
 :
 : SEQ ID NO 16593
 :
 : LENGTH: 308
 : TYPE: prt
 :
 : ORGANISM: Candida albicans
 :
 : US-09-248-796A-16593

Query Match	26.3%	Score 372.5;	DB 4;	Length 308;
Best Local Similarity	36.2%	Pred. No. 1.7e-32;		
Matches 97, Conservative	47;	Mismatches 105;	Indels 19;	Gaps 8;

```

Qy      5  I P P E R D G H I L T Y G A C G G L A E A L I K G L L A A G S D I A L D I D O E K T A A Q A E H Y H A T E E L   64
Db      56  L D A F S L K G V A A Y T G S S G G I G W A V A E G I A Q A D A D A I M - - I N S H P A D D K A E Y - - - - - L   106

Qy      65  K L A E V P M G S Y A C D I S D S D T V A K V P A Q A O F G K L P L H V N T A G C E N F P C E D Y P A Q A E   124
Db      107 T K T G Y V K S R A K Y C N V D P O D V E K V V Q I E L D E G T I D I F V A A G A V M E G F - - E I D V G V D   164

Qy      125  K W K V - - N I L G S I Y V S Q A F A K P L I K E G I K G A S V V I I G S M G A I V N D P O N Q V V N M S R A G V   182
Db      165  K W N V D V D L A N S Y T Y C A H V V G P I F R K K G G - S F I T A S M S I V N V P Q L Q A A V A A K A G V   223

Qy      183  I H L A K T L A C E W A K Y N I R V N S L N F G Y T G B L Y K N V I N G S E B E L Y N R W I S G I P O O R S S E K E Y   242
Db      224  K H L S K S I S V M A E P - A R V N S V S P G Y I A T H I A S E - - - F A D P D V K S K W L Q U T P L G R E A K P R E L   279

Qy      243  I G A V I Y L I S E S A S Y T T G A S I L A V D G G E T   270
Db      280  V G A I Y L I A S D - A A S Y T T G A D I A V D G G Y T   306

```

```

RESULT 8
US-09-489-039A-12990
; Sequence 12990, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489, 039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12990
; LENGTH: 316
; TYPE: PRF
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12990

```

[illegible]

Oy 184 HLAQTACCAKAKNIRVNSLNGYLYGTLTRNVINGBELYNR---MISGLPOORSEPK 240
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Dd 231 HLKSLLAMWVGKGIFRVNSISPGYTATPM-----NTREPMVHQREFESQTPMRMAKVE 285

Oy 241 EYICAVILYLISEAASVTTGASILDGGFTSM 272
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Dd 286 ELNGPALFLASD-NAHFCTGYDLVDVGGFVCM 316

```

RESULT 9
US-09-134-001C-3562
; Sequence 3562, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3562
; LENGTH: 257
; TYPE: prt
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3562

```

Query Match	24.1%	Score 340.5	DB 3	Length 257
Best Local Similarity	29.9%	Pred No. 4.4e-29		
Matches	82	Conservative 64	Mismatches 105	Indels 23
				Gaps 6
Qy	1	MTDIYPTFRFDGHLTYTTCGCGGLAEALIKGLHALYSGDIALDLIDQ--EKTAKOAEYH	57	
Db	3	MNNINIEKPLTGTGKAMVNIIGATIGICKAAAEALAQGANIVLADLSNIGQETATISTOS	62	
Qy	58	KYATEELKIKVEYPKKMGYSACDISDSOTYHAKPFAQYAKDFGKLPHLVNTAGCEMPCED	117	
Db	63	GVKXTSLKL-----DITHHDEVQIADYVVRBEGKIDI-LVNNASISIDDTTEN	110	
Qy	118	YPAKNAEKKNVKNLGLSLYSQAFAPLKIKSIIKGSAYVLISMSGAYVNDPONOVVNM	177	
Db	111	ISYEWLTKELINSINGAFSVAQTVGRQMEKKA--SGSMITVSSVGLIANKTODOSYTT	168	
Qy	178	SKAGVTHLAKTLACCEKAKYNIRVNSLNPXYIGPLTKVNIINSNEELIYNRMISGIDPOORMS	237	
Db	169	SKAGVMTLTKSLAREMSRYGIKINAIAPISQYMTIETIEKLIANDT-----TMTTPHERVG	223	
Qy	238	EPKEVYIGAVLYLISBSGAASTTGASLLVDGGETS	271	
Db	224	EPBELAGITVYTLASD--ASSFTQSGSVNIDIGSYSA	256	

```

RESULT 10
US-09-252-991A-30474
; Sequence 30474, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,180
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30474

```

LENGTH: 256
 TYPE: PRT
 ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-30474

Query Match 22.8%; Score 322.5; DB 4; Length 256;
 Best Local Similarity 34.7%; Pred. No. 4,3e-27;
 Matches 93; Conservative 48; Mismatches 104; Indels 23; Gaps 10;

8 FRFDGLTIVTGAACGGLAEALIKGLIAYGSDIALDIDOEKTAQAQAEYHKYATELKLK 67
 8 FDLOSKIYFVGASGIGIEALIKLAQGAHYI--VSRKIDGQ-----AVADALTA 58
 QY 68 EVPKKGSYACDISDSTVHKYFAQVAKFGKLPRLVNTAG---YCENFPCEDYPAKXA 123
 DB 59 ECGKATATACHIGEWEQIQWFAQIREQFGRLDI-LVNNAAATNPQFC-NVLETLDGA--F 114
 QY 124 EKMVNVNLSGLYVQAFAKPLIKGIGASVVLIGSGAIVNDPQOVVYNNKAGVY 183
 DB 115 OKTVDNIRGYFMSIEGSKLMKEHG--GGSIINVASING--VSPGEFGIYSVTKAAYI 170
 QY 184 HLAATLACEMAKYINRVNSLNPGYIYGPITKXVINGNEELVYRWISGIPQORMSEPKYI 243
 DB 171 SMTVFAECQAQFGRCNALIPGLDTKFAALYK-NDAIRNLALQRIPLKVAEPESEMA 229
 QY 244 GAVLYLSESAASYTTGASLLVDGGFTS 271
 DB 230 GAVLYLASE-ASSYTTGVALNVDGGFLS 256

RESULT 11
 US-09-363-189B-6
 Sequence 6, Application US/09363189B
 Patent No. 6242228
 GENERAL INFORMATION:

APPLICANT: SUGIYAMA, MASAKAZU
 APPLICANT: TONOUCHI, NAOTO
 APPLICANT: SUZUKI, SHUNICHI
 APPLICANT: YOKOZAKI, KENZO
 TITLE OF INVENTION: XYLITOL DEHYDROGENASE OF ACETIC ACID BACTERIA AND GENE THEREOF
 FILE REFERENCE: 0010-1024-0
 CURRENT APPLICATION NUMBER: US/09/363,189B
 CURRENT FILING DATE: 1999-07-26
 PRIOR APPLICATION NUMBER: JP10-216047
 PRIOR FILING DATE: 1998-07-30
 NUMBER OF SEQ ID NOS: 16
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 6
 LENGTH: 262
 TYPE: PRT
 ORGANISM: Gluconobacter oxydans
 US-09-363-189B-6

Query Match 22.7%; Score 321.5; DB 3; Length 262;
 Best Local Similarity 31.0%; Pred. No. 5.8e-27;
 Matches 86; Conservative 53; Mismatches 103; Indels 35; Gaps 10;

9 RFDGLTIVTGAACGGLAEALIKGLIAYGSDIALDIDOEKTAQAQAEYHKYATELKLK 68
 4 KFNKGVCLVTAGAGNIGLATYLRLEBETATLALDNKNLEAKABAYREKGV----- 57
 QY 69 VPKKGSYACDISDSTVHKYFAQVAKFGKLPRLVNTAGYCENF-PCEDYPAKXAEKAV 127
 DB 58 ---ARSTYCDVTSEBAVIGTVDSVVRDGRKIDF-LFNNAAGYOGAFAPQODIPSDPRAVL 113
 QY 128 KVNLSGLYVQAFAKPLIKGIGASVVLIGSGAIVNDPQOVVYNNKAGVYHILAK 187
 DB 114 TIVTGAFAHVAKASRQITQNY--GRIVNTASVAG--VKBPYMAAVAGSKGAILALTE 169
 QY 188 TLACEMAKYINRVNSLNPGYIYGP-----LTKVYINGNE-----ELVYRWISGI 231
 DB 170 TVALDLAPYINRVNALSFGH--GGFPMKEROVELOAKV--GSQYFSTDPKVAQOQIGSV 226

QY 232 PQORMSEPKYIYAVLYLSESAASYTTGASLLVDGG 268
 DB 227 PMRYVDINEIPGVAFILGDD-SSFMTGVNLPIAGG 262

RESULT 12
 US-09-468-738A-2
 Sequence 2, Application US/09468738A
 Patent No. 6312933
 GENERAL INFORMATION:

APPLICANT: Kimoto, No. 6312933hiro
 APPLICANT: Yamamoto, Hiroaki
 APPLICANT: Mitsuhashi, Kazuya
 TITLE OF INVENTION: NOVEL CARBONYL REDUCTASE, METHOD FOR PRODUCING SAID ENZYME, DNA
 TITLE OF INVENTION: ENCODING SAID ENZYME, AND METHOD FOR PRODUCING ALCOHOL USING SAID
 TITLE OF INVENTION: ENZYME
 FILE REFERENCE: 06501-050001
 CURRENT APPLICATION NUMBER: US/09/468,738A
 CURRENT FILING DATE: 1999-12-21
 PRIOR APPLICATION NUMBER: JP 1999-171160
 PRIOR FILING DATE: 1999-06-17
 PRIOR APPLICATION NUMBER: JP 1998-363130
 NUMBER OF SEQ ID NOS: 29
 SOFTWARE: PatentIn Ver. 2.0, reformatted using WordPerfect 5.1
 SEQ ID NO 2
 LENGTH: 292
 TYPE: PRT
 ORGANISM: Kluyveromyces aestuarii
 US-09-468-738A-2

Query Match 22.1%; Score 312.5; DB 3; Length 292;
 Best Local Similarity 32.3%; Pred. No. 6.9e-26;
 Matches 93; Conservative 50; Mismatches 98; Indels 47; Gaps 10;

3 DYITFRPDGLTIVTGAACGGLAEALIKGLIAYGSDIALDIDOEKTAQAQAEYHKYATE 62
 30 NYLSLFGQKGLTIVTGAAGALGALCEGFASCGSDVVIID-----YKTSPE 76
 QY 63 ---ELKKEVYKGSYACDISDSTVHKYFAQVAKFGKLPRLVNT---AGYC----- 110
 DB 77 LGSVLESRYGVRSKYVDITSSSEVDLVVAKLIEDP---PDRINFVANAIAWNTGS 133
 QY 111 ---ENFPCEDYPAKXAEKAVYVNLGSLYVQAFAKPLIKGIGASVVLIGSGAIVN 167
 DB 134 IINENATPDVW-----KRVVDVYQGTTHCAKYAEYFKQOG--HGVILFASSTYSIN 186
 QY 168 DPQOVVYNNKAGVYHILATLACEMAKY-----NIRVNSLNPGYIYGPITKXVINGNEE 222
 DB 187 VPNYQTCYNASKAIVRHMAKFAVEFAHLTNPAGKIRCNSVSPGYTDALSAFY---PVE 243
 QY 223 LYRWISGIPQORMSEPKYIYAVLYLSESAASYTTGASLLVDGGFT 270
 DB 244 QRAQWGLTTPWGREALPQELVGAITYLASD-AASFNGCDIQVDGGYT 290

RESULT 13
 US-09-940-019-2
 Sequence 2, Application US/09940019
 Patent No. 6416986
 GENERAL INFORMATION:

APPLICANT: Kimoto, No. 6416986hiro
 APPLICANT: Yamamoto, Hiroaki
 APPLICANT: Mitsuhashi, Kazuya
 TITLE OF INVENTION: NOVEL CARBONYL REDUCTASE, METHOD FOR PRODUCING SAID ENZYME, DNA
 TITLE OF INVENTION: ENCODING SAID ENZYME, AND METHOD FOR PRODUCING ALCOHOL USING SAID
 FILE REFERENCE: 06501-050001
 CURRENT APPLICATION NUMBER: US/09/940,019
 CURRENT FILING DATE: 2001-08-27
 PRIOR APPLICATION NUMBER: 09/468,738
 PRIOR FILING DATE: 1999-06-17
 PRIOR APPLICATION NUMBER: JP 1998-363130

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 10:31:08 ; Search time 162 Seconds

(without alignments)
648,653 Million cell updates/sec

Title: US-10-720-018-2

Perfect score: 1415
Sequence: 1 MTDIYPTFRFGHLLTVTGA.....SAASYTTGASLLVDGPTSW 272

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1726216 seqs, 386330316 residues

Total number of hits satisfying chosen parameters: 1726216

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA.*
1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubppaa/US11_PUBCOMB.pep.*
20: /cgn2_6/ptodata/2/pubppaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1415	100.0	272	US-10-720-018-2	Sequence 2, Appl1
2	402.5	28.4	257	US-10-369-493-2923	Sequence 2923, Ap
3	396	28.0	211	US-10-369-493-13036	Sequence 13036, A
4	381	26.9	283	US-10-639-159-42	Sequence 42, Appl
5	381	26.9	283	US-10-639-159-60	Sequence 60, Appl
6	381	26.9	283	US-10-782-258-42	Sequence 42, Appl
7	381	26.9	283	US-10-782-258-60	Sequence 60, Appl
8	380	26.9	283	US-10-639-159-40	Sequence 40, Appl
9	380	26.9	283	US-10-782-258-40	Sequence 40, Appl
10	379	26.8	283	US-10-782-258-118	Sequence 118, Ap
11	378	26.7	283	US-09-777-157A-1	Sequence 1, Appl1

12	378	26.7	283	10	US-09-734-237B-70	Sequence 70, Appl
13	378	26.7	283	16	US-10-639-159-2	Sequence 2, Appl1
14	378	26.7	283	16	US-10-639-159-6	Sequence 6, Appl1
15	378	26.7	283	16	US-10-639-159-8	Sequence 8, Appl1
16	378	26.7	283	16	US-10-782-258-2	Sequence 2, Appl1
17	378	26.7	283	16	US-10-782-258-6	Sequence 6, Appl1
18	378	26.7	283	16	US-10-782-258-8	Sequence 8, Appl1
19	376	26.6	253	15	US-10-369-493-11548	Sequence 11548, A
20	376	26.6	283	16	US-10-639-159-46	Sequence 46, Appl
21	376	26.6	283	16	US-10-639-159-54	Sequence 54, Appl
22	376	26.6	283	16	US-10-782-258-4	Sequence 4, Appl1
23	376	26.6	283	16	US-10-782-258-46	Sequence 46, Appl
24	376	26.6	283	16	US-10-782-258-54	Sequence 54, Appl
25	376	26.6	283	16	US-10-782-258-116	Sequence 116, App
26	375	26.5	283	16	US-10-782-258-120	Sequence 120, App
27	375	26.5	283	15	US-10-369-493-1304	Sequence 3304, Ap
28	374	26.4	245	15	US-10-639-159-44	Sequence 44, Appl
29	374	26.4	283	16	US-10-782-258-44	Sequence 44, Appl
30	374	26.4	283	16	US-10-639-159-50	Sequence 50, Appl
31	373	26.4	283	16	US-10-639-159-56	Sequence 56, Appl
32	373	26.4	283	16	US-10-639-159-58	Sequence 58, Appl
33	373	26.4	283	16	US-10-782-258-56	Sequence 56, Appl
34	373	26.4	283	16	US-10-782-258-50	Sequence 50, Appl
35	373	26.4	283	16	US-10-782-258-58	Sequence 58, Appl
36	373	26.4	283	16	US-10-782-258-114	Sequence 114, App
37	373	26.4	283	16	US-10-639-159-52	Sequence 52, Appl
38	370	26.1	283	16	US-10-782-258-52	Sequence 52, Appl
39	370	26.1	283	15	US-10-369-493-14958	Sequence 14958, A
40	367	25.9	251	15	US-10-639-159-48	Sequence 48, Appl
41	367	25.9	283	16	US-10-782-258-48	Sequence 48, Appl
42	367	25.9	263	14	US-10-128-714-8283	Sequence 8283, Ap
43	364	25.7	251	15	US-10-369-493-14214	Sequence 14214, A
44	362	25.6	251	15	US-10-508-681-3	Sequence 3, Appl1
45	358.5	25.3	266	18		

ALIGNMENTS

RESULT 1
US-10-720-018-2
; Sequence 2, Application US/10720018
; Publication No. US20040132074A1
; GENERAL INFORMATION:
; APPLICANT: Verbio, Rltva
; APPLICANT: Richard, Peter
; APPLICANT: Penttila, Merja
; TITLE OF INVENTION: New Enzyme for an in vivo and in vitro utilisation of
; FILE REFERENCE: 2530-120
; CURRENT APPLICATION NUMBER: US/10/720,018
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Ambrosiozyma monospora
US-10-720-018-2

Query Match 100.0%; Score 1415; DB 16; Length 272;
Best Local Similarity 100.0%; Pred. No. 5,4e-131;
Matches 272; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTDIYPTFRFGHLLTVTGAAGGAGLALIKGLAYGSDIALDDIDDKTAQAQAEYHRYA 60
DB 1 MTDIYPTFRFGHLLTVTGAAGGAGLALIKGLAYGSDIALDDIDDKTAQAQAEYHRYA 60
QY 61 TEELKLEVPKMGVYADIGSDPTVHVPAPVADBEKPLHLVNTAGYCNFCEYPA 120
DB 61 TEELKLEVPKMGVYADIGSDPTVHVPAPVADBEKPLHLVNTAGYCNFCEYPA 120
QY 121 KNAEMKVNLLAGSLVYGAFAKPLIKGIGKASVVLIGSWGAIIVNDPQNVVYNNNSKA 180

Db 121 KNAERKVVNLLGSLYVSOAFKPLIKGIGASVVLIGMSGALVNDPQVQVYNNMSKA 180
Qy 181 GVIIHLAKTLACEMAKNIRVNSLNPGYIYGPLTKRVINGNEELNRRYISGIPQORMSPK 240
Db 181 GVIIHLAKTLACEMAKNIRVNSLNPGYIYGPLTKRVINGNEELNRRYISGIPQORMSPK 240
Qy 241 EYIGAVLYLSESAASYTTGASLVDGGFTSM 272
Db 241 EYIGAVLYLSESAASYTTGASLVDGGFTSM 272

RESULT 2
US-10-369-493-2923
Sequence 2923, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xiandeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 2923
LENGTH: 257
TYPE: PRT
ORGANISM: Thermotoga maritima
US-10-369-493-2923

Query Match 28.4%; Score 402.5; DB 15; Length 257;
Best Local Similarity 36.9%; Pred. No. 5,1e-31;
Matches 101; Conservative 50; Mismatches 90; Indels 33; Gaps 11;

Qy 8 FRPDGHLITVTGACGGALBALIKGLIAYGSDIALDIDDEKTAQAQAEYHKYATEBLKIK 67
Db 8 FSLKRKVALVTGGGQIGIKALQALAAAGAAVLIMDINEE-TARRIVE----- 54
Qy 68 EYIPKMS---YACDSDSDYTHAKVQAQAKDPRGLPLVNTAG---YCENPCEPYA 120
Db 55 EIKKGGKADFYVDYTKEDCEGAVKALDRWKLIDIG-VNNAIGDWCE--AENYFV 110
Qy 121 KNAERKVVNLLGSLYVSOAFKPLIKGIGASVVLIGMSGALVNDPQVQVYNNMSKA 180
Db 111 EKMKKVIDNLYG-VFLSKAEFHAMKER-KYCKIINIMSMGHIYNNKQKQTAAYASKA 168
Qy 181 GVIIHLAKTLACEMAKNIRVNSLNPGYIYGPL-TKNVINGNEELNRRYISGIPQORMSPK 238
Db 169 GVIIHLRSLAAEWAPYRIVNSISPGYIRPLIESBNV---KDLVPLWMLPIGLRGE 224
Qy 239 PKEYIGAVLYLSESAASYTTGASLVDGGFTSM 272
Db 225 VDDLIGAAIFLAS-PASDYVTGHDVLVDGGYTW 257

RESULT 3
US-10-369-493-13036
Sequence 13036, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xiandeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 13036
LENGTH: 211
TYPE: PRT
ORGANISM: Aspergillus nidulans
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(211)
OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-13036

FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 13036
LENGTH: 211
TYPE: PRT
ORGANISM: Aspergillus nidulans
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(211)
OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-13036

Query Match 28.0%; Score 396; DB 15; Length 211;
Best Local Similarity 41.1%; Pred. No. 1.7e-30;
Matches 88; Conservative 38; Mismatches 64; Indels 24; Gaps 4;
Qy 78 DISDSDYTHAKVFAQVAKDPGLPLHLVNTAGYCENPCEDEYPAKNAERKVVNLLGSLYV 137
Db 1 DVSNDPVSNNALSEVIKHKID-HLVTSAGFTENFPAISYPRMQLMGVNDGYTLF 59
Qy 138 SOAFKPLIKGIGASVVLIGMSGALVNDPQVQVYNNMSKGVIIHLAKTLACEMAKN 197
Db 60 ATGAKHLMERKADG-SIVMIGMSGALVNPQQAAYNAKAVRHLASSFAEWAGHG 118
Qy 198 IRVNSLNPGYIYGPL-----TKNVINGNEELNRRYISGIPQORM 236
Db 119 IRVNCISPGVWLTALXYPESIQNSXARTICKFVATRRKILIDENPELRDKMISILPFGKM 178
Qy 237 SEPKYIGAVLYLSESAASYTTGASLVDGGFT 270
Db 179 GVPEDMGAVTFILSD-ASKYITGADLRVDGYT 211

RESULT 4
US-10-639-159-42
Sequence 42, Application US/10639159
Publication No. US20040137585A1
GENERAL INFORMATION:
APPLICANT: Davis, S. Christopher
APPLICANT: Grate, John H.
APPLICANT: Gray, David R.
APPLICANT: Gruber, John M.
APPLICANT: Huiteman, Gjalte W.
APPLICANT: Ma, Steven K.
APPLICANT: Newman, Lisa M.
APPLICANT: Sheldon, Roger
APPLICANT: Wang, Li A
TITLE OF INVENTION: ENZYMAIC PROCESSES FOR THE PRODUCTION
FILE REFERENCE: 0339.210US
CURRENT APPLICATION NUMBER: US/10/639,159
CURRENT FILING DATE: 2003-08-11
PRIOR APPLICATION NUMBER: US 60/402,436
PRIOR FILING DATE: 2002-08-09
NUMBER OF SEQ ID NOS: 92
SOFTWARE: FaatSEO for Windows Version 4.0
SEQ ID NO 42
LENGTH: 283
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: KRED krt215
US-10-639-159-42

Query Match 26.9%; Score 381; DB 16; Length 283;
Best Local Similarity 36.1%; Pred. No. 7.9e-29;
Matches 97; Conservative 52; Mismatches 96; Indels 22; Gaps 8;
Qy 8 FRPDGHLITVTGACGGALBALIKGLIAYGSDIALDIDDEKTAQAQAEYHKYATEBLKIK 67

```

Db      29  FRLNGKVASITSSSSGIGALAEAFQVADVAITNSQDATGKAEALAKKTCV----- 82
      68  EVPRKGSYACDISDSTVHKFAQVAKDFGKLP.LHLVNTAG--YCENFPCEDYPAKNAEK 125
      83  ---KVKAYKANVSSSDAVKQTEIQIKDFGHLDIVAN-AGIPWKGAVIDDDDKHFDQ 138
Qy      126  MKVKNLLGSLYVS---QAFAPLKEGIGKASVVLIGSGAIVNDPONOVVNNMSKAG 181
      139  VIDDLKGQVYAKHAGHRRERFEKIGKA-LIFTSVSGHIVNIPOFQATYNAAKAG 197
Db      182  VHLAKTLACEMAKYINRVNSLNPGYIGPLTKVNINGNEELYNRMISGIPQRMSEPRE 241
      198  VRHFAKSLAVEFAPF-ARVNSVSPGYINTERTSDFV---PQETQKMWSLVPLGRGGETAE 253
Qy      242  YIGAVLYLSEBSAASYTTGASLVDGGFT 270
      254  LVGAVLFLASD-AGSYATGTDTIIVDGGYT 281

```

RESULT 5
US-10-639-159-60

```

; Sequence 60, Application US/10639159
; Publication No. US20040137585A1
; GENERAL INFORMATION:
; APPLICANT: Davis, S. Christopher
; APPLICANT: Grate, John H.
; APPLICANT: Gray, David R.
; APPLICANT: Gruber, John M.
; APPLICANT: Hulsman, Gjalte W.
; APPLICANT: Ma, Steven K.
; APPLICANT: Newman, Lisa M.
; APPLICANT: Sheldon, Roger
; APPLICANT: Wang, Li A.
; TITLE OF INVENTION: ENZYMATIC PROCESSES FOR THE PRODUCTION
; TITLE OF INVENTION: OF 4-SUBSTITUTED-3-HYDROXYBUTYRIC ACID DERIVATIVES
; FILE REFERENCE: 0339.210US
; CURRENT APPLICATION NUMBER: US/10/639.159
; PRIOR FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: US 60/402,436
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: KRED ktrh95
; US-10-639-159-60

```

Query Match 26.9%; Score 381; DB 16; Length 283;
Best Local Similarity 36.1%; Pred. No. 7.9e-29;
Matches 97; Conservative 52; Mismatches 96; Indels 22; Gaps 8;

```

Qy      8  FRPDGHLITVAGACGALAEALIKGLAYGSDIALDIDQEKTAQAQAEHYKATBELTK 67
      29  FKLNGKVASITSSSSGIGYALAEAFQVADVAITWNSQDATGKAEALAKKTCV----- 82
      68  EVPRKGSYACDISDSTVHKFAQVAKDFGKLP.LHLVNTAG--YCENFPCEDYPAKNAEK 125
      83  ---KVKAYKANVSSSDAVKQTEIQIKDFGHLDIVAN-AGIPWKGAVIDDDDKHFDQ 138
Db      126  MKVKNLLGSLYVS---QAFAPLKEGIGKASVVLIGSGAIVNDPONOVVNNMSKAG 181
      139  VIDDLKGQVYAKHAGHRRERFEKIGKA-LIFTSVSGHIVNIPOFQATYNAAKAG 197
Qy      182  VHLAKTLACEMAKYINRVNSLNPGYIGPLTKVNINGNEELYNRMISGIPQRMSEPRE 241
      198  VRHFAKSLAVEFAPF-ARVNSVSPGYINTERTSDFV---PQETQKMWSLVPLGRGGETAE 253
Db      242  YIGAVLYLSEBSAASYTTGASLVDGGFT 270
      254  LVGAVLFLASD-AGSYATGTDTIIVDGGYT 281

```

Db 254 LVGAVLFLASD-AGSYATGTDTIIVDGGYT 281

RESULT 6
US-10-782-258-42

```

; Sequence 42, Application US/10782258
; Publication No. US20040214297A1
; GENERAL INFORMATION:
; APPLICANT: Davis, S. Christopher
; APPLICANT: Grate, John H.
; APPLICANT: Gray, David R.
; APPLICANT: Gruber, John M.
; APPLICANT: Hulsman, Gjalte W.
; APPLICANT: Ma, Steven K.
; APPLICANT: Newman, Lisa M.
; APPLICANT: Sheldon, Roger
; APPLICANT: Wang, Li A.
; TITLE OF INVENTION: Enzymatic Processes for the Production
; TITLE OF INVENTION: of 4-Substituted 3-Hydroxybutyric Acid Derivatives and
; FILE REFERENCE: 0339.310US
; CURRENT APPLICATION NUMBER: US/10/782.258
; PRIOR FILING DATE: 2004-02-18
; PRIOR APPLICATION NUMBER: US 10/639,159
; PRIOR FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: US 60/402,436
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: KRED ktrh215
; US-10-782-258-42

```

Query Match 26.9%; Score 381; DB 16; Length 283;
Best Local Similarity 36.1%; Pred. No. 7.9e-29;
Matches 97; Conservative 52; Mismatches 96; Indels 22; Gaps 8;

```

Qy      8  FRPDGHLITVAGACGALAEALIKGLAYGSDIALDIDQEKTAQAQAEHYKATBELTK 67
      29  FKLNGKVASITSSSSGIGYALAEAFQVADVAITWNSQDATGKAEALAKKTCV----- 82
      68  EVPRKGSYACDISDSTVHKFAQVAKDFGKLP.LHLVNTAG--YCENFPCEDYPAKNAEK 125
      83  ---KVKAYKANVSSSDAVKQTEIQIKDFGHLDIVAN-AGIPWKGAVIDDDDKHFDQ 138
Qy      126  MKVKNLLGSLYVS---QAFAPLKEGIGKASVVLIGSGAIVNDPONOVVNNMSKAG 181
      139  VIDDLKGQVYAKHAGHRRERFEKIGKA-LIFTSVSGHIVNIPOFQATYNAAKAG 197
Db      182  VHLAKTLACEMAKYINRVNSLNPGYIGPLTKVNINGNEELYNRMISGIPQRMSEPRE 241
      198  VRHFAKSLAVEFAPF-ARVNSVSPGYINTERTSDFV---PQETQKMWSLVPLGRGGETAE 253
Qy      242  YIGAVLYLSEBSAASYTTGASLVDGGFT 270
      254  LVGAVLFLASD-AGSYATGTDTIIVDGGYT 281

```

RESULT 7

```

; Sequence 60, Application US/10782258
; Publication No. US20040214297A1
; GENERAL INFORMATION:
; APPLICANT: Davis, S. Christopher
; APPLICANT: Grate, John H.
; APPLICANT: Gray, David R.
; APPLICANT: Gruber, John M.
; APPLICANT: Hulsman, Gjalte W.
; APPLICANT: Ma, Steven K.

```


QY 68 EVPMKGSYACDISDPTVHKVFAQVAKDFGLPLHLVNTAG--YCNFPCEDYPAKNAEK 125
 DB 83 ---KVAIVKANVSSSDAVAKQTEEQIKDGHLDIVAN-AGIPMTKGAVIIDDDDKHFDQ 138
 QY 126 MVKNLGLSLYVS---QAFAPLKEGIGKASVVLIGSMGAIYNDPONOVVNMKAG 181
 DB 139 VVDVLDKGVGYAKAGHRRERPEKEGKGA-LVFTASMSGHIVNVPFOATTNAKAG 197
 QY 182 VHLAKTLACEMAKNIRVNSLNPGYITPLTKNVINGNEELYNRWISGIPQORMSEPK 241
 DB 198 VRHFAKSLAVEFAPF-ARVNSVSPGYINTBISDFV---PQETQKMWSLVPLGRGETAE 253
 QY 242 YIGAVLYLSSASAYTTGASLVDGGET 270
 DB 254 LVGATYFLASD-AGSYATGTDTIIDVGGYT 281

RESULT 10
 US-10-782-258-118
 ; Sequence 118, Application US/10782258
 ; Publication No. US20040214297A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Davis, S. Christopher
 ; APPLICANT: Grate, John H.
 ; APPLICANT: Gruber, John M.
 ; APPLICANT: Hultman, Gjalte W.
 ; APPLICANT: Ma, Steven K.
 ; APPLICANT: Newman, Lisa M.
 ; APPLICANT: Sheldon, Roger
 ; APPLICANT: Wang, Li A.
 ; TITLE OF INVENTION: Enzymatic Processes for the Production
 ; TITLE OF INVENTION: of 4-Substituted 3-Hydroxybutyric Acid Derivatives and
 ; FILE REFERENCE: 0339.310US
 ; CURRENT APPLICATION NUMBER: US/10/782,258
 ; PRIOR FILING DATE: 2004-02-18
 ; PRIOR APPLICATION NUMBER: US 10/639,159
 ; PRIOR FILING DATE: 2003-08-11
 ; PRIOR APPLICATION NUMBER: US 60/402,436
 ; PRIOR FILING DATE: 2002-08-09
 ; NUMBER OF SEQ ID NOS: 126
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 118
 ; LENGTH: 283
 ; TYPE: PRF
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: KRED S01091625
 ; US-10-782-258-118

Query Match 26.8%; Score 379; DB 16; Length 283;
 Best Local Similarity 37.2%; Pred. No. 1.2e-28;
 Matches 100; Conservative 47; Mismatches 100; Indels 22; Gaps 8;
 QY 8 FRPDGLTIVTGACGGLAEALIKGLAYGSDIALIDIOEKTAKQAQEHYKATEELTK 67
 DB 29 FKLNGKVASITSGNSGIGYALAEAPAGADVAIWMNSHDATGKAEALAKKYGV----- 82
 QY 68 EVPMKGSYACDISDPTVHKVFAQVAKDFGLPLHLVNTAG--YCNFPCEDYPAKNAEK 125
 DB 83 ---KVAIVKANVSSSDAVAKQTEEQIKDGHLDIVAN-AGIPMTKGAVIIDDDDKHFDQ 138
 QY 126 MVKNLGLSLYVS---QAFAPLKEGIGKASVVLIGSMGAIYNDPONOVVNMKAG 181
 DB 139 VVDVLDKGVGYAKAGHRRERPEKEGKGA-LVFTASMSGHIVNVPFOATTNAKAG 197
 QY 182 VHLAKTLACEMAKNIRVNSLNPGYITPLTKNVINGNEELYNRWISGIPQORMSEPK 241
 DB 198 VRHFAKSLAVEFAPF-ARVNSVSPGYINTBISDFV---PQETQKMWSLVPLGRGETAE 253
 QY 242 YIGAVLYLSSASAYTTGASLVDGGET 270
 DB 254 LVGATYFLASD-AGSYATGTDTIIDVGGYT 281

DB 254 LVGATYFLASD-AGSYATGTDTIIDVGGYT 281
 RESULT 11
 US-09-777-157A-1
 ; Sequence 1, Application US/09777157A
 ; Publication No. US2002000651A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Yasohara, Yoshihiko
 ; APPLICANT: Kizaki, No. 6448052iyuki
 ; APPLICANT: Hasegawa, Junzo
 ; APPLICANT: Mada, Masaru
 ; APPLICANT: Shimizu, Sakayu
 ; APPLICANT: Katoka, Michihiko
 ; APPLICANT: Yamamoto, Kazuhiko
 ; APPLICANT: Kawabata, Hiroshi
 ; APPLICANT: Kita, Keiko
 ; TITLE OF INVENTION: Carbonyl Reductase Enzyme and Methods for its Use
 ; FILE REFERENCE: 068383.0110
 ; CURRENT APPLICATION NUMBER: US/09/777,157A
 ; PRIOR FILING DATE: 2001-02-05
 ; PRIOR APPLICATION NUMBER: US 09/367,012
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: Patent version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 283
 ; TYPE: PRF
 ; ORGANISM: Candida magnoliae
 ; US-09-777-157A-1

Query Match 26.7%; Score 378; DB 9; Length 283;
 Best Local Similarity 36.8%; Pred. No. 1.6e-28;
 Matches 99; Conservative 48; Mismatches 100; Indels 22; Gaps 8;
 QY 8 FRPDGLTIVTGACGGLAEALIKGLAYGSDIALIDIOEKTAKQAQEHYKATEELTK 67
 DB 29 FKLNGKVASITSGNSGIGYALAEAPAGADVAIWMNSHDATGKAEALAKKYGV----- 82
 QY 68 EVPMKGSYACDISDPTVHKVFAQVAKDFGLPLHLVNTAG--YCNFPCEDYPAKNAEK 125
 DB 83 ---KVAIVKANVSSSDAVAKQTEEQIKDGHLDIVAN-AGIPMTKGAVIIDDDDKHFDQ 138
 QY 126 MVKNLGLSLYVS---QAFAPLKEGIGKASVVLIGSMGAIYNDPONOVVNMKAG 181
 DB 139 VVDVLDKGVGYAKAGHRRERPEKEGKGA-LVFTASMSGHIVNVPFOATTNAKAG 197
 QY 182 VHLAKTLACEMAKNIRVNSLNPGYITPLTKNVINGNEELYNRWISGIPQORMSEPK 241
 DB 198 VRHFAKSLAVEFAPF-ARVNSVSPGYINTBISDFV---PQETQKMWSLVPLGRGETAE 253
 QY 242 YIGAVLYLSSASAYTTGASLVDGGET 270
 DB 254 LVGATYFLASD-AGSYATGTDTIIDVGGYT 281

RESULT 12
 US-09-734-237B-70
 ; Sequence 70, Application US/09734237B
 ; Publication No. US20030064432A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rozzelli, J. David
 ; APPLICANT: Bui, Peter
 ; APPLICANT: Hua, Ling
 ; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
 ; FILE REFERENCE: B583.40608
 ; CURRENT APPLICATION NUMBER: US/09/734,237B
 ; PRIOR FILING DATE: 2000-12-08
 ; PRIOR APPLICATION NUMBER: 09/494,921
 ; NUMBER OF SEQ ID NOS: 79
 ; SOFTWARE: Patent version 3.1
 ; SEQ ID NO 70

```

; LENGTH: 283
; TYPE: PRT
; ORGANISM: Candida magnoliae
US-09-734-237B-70

Query Match          26.7%; Score 378; DB 10; Length 283;
Best Local Similarity 36.8%; Pred. No. 1,6e-28;
Matches 99; Conservative 48; Mismatches 100; Indels 22; Gaps 8;

QY      8 FRPDGHLTVTVGACGGALAEALIKGLLAYGSDIALIDDOEKTAAKQAEYHKYATEELKXK 67
DB      29 FLKNGKVASITGSSSGIGYALAEAPQAGDAVALIWNSHDATGKAELATAKYGIV----- 82
QY      68 EYPRKGSYACDISDSDYTHKVFPAQVAKDPKPLHLVNTAG--YCENFPCEDYPAKNAEK 125
DB      83 ---KPKAKVANSSSDAVKQTIEQIKPFGHLDIVAN-AGIPWTKGAYIDODDDKHPDQ 138
QY      126 MYKVVLLLSLYS----QAFAPKLKEGIKGASVVLIGSMGSAIVNDPONOVYNNMSKXG 181
DB      139 VVDVDLKGVGIVAKHAGHFRFRFEKKEKKA-LVFTASMSGHIVNVPQFQATYNAAKAG 197
QY      182 VYHLAKTLACEWAKXNIRVNSINPGYIVGLTKNVINGNEELVNMISGIPQQRMSPEKE 241
DB      198 VRHFAKSLAVEPAP-ARVNSVSPGYINTIEDSFV---PQETQRKMSVLVPLGRGGETAE 253
QY      242 YTGAVLYLISESAAYTTGASILLVNGGFT 270
DB      254 LVGAYLFLASD-AGSYATGTDTIIVDGGYT 281

RESULT 13
US-10-639-159-2
; Sequence 2, Application US/10639159
; Publication No. US20040137585A1
; GENERAL INFORMATION:
; APPLICANT: Davis, S. Christopher
; APPLICANT: Grate, John H.
; APPLICANT: Gray, David M.
; APPLICANT: Gruber, John M.
; APPLICANT: Huisman, Gjalte W.
; APPLICANT: Ma, Steven K.
; APPLICANT: Newman, Lisa M.
; APPLICANT: Sheldon, Roger
; APPLICANT: Wang, Li A
; TITLE OF INVENTION: ENZYMATIC PROCESSES FOR THE PRODUCTION
; TITLE OF INVENTION: OF 4-SUBSTITUTED-3-HYDROXYBUTYRIC ACID DERIVATIVES
; FILE REFERENCE: 0339.210US
; CURRENT APPLICATION NUMBER: US/10/639, 159
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: US 60/402,436
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FaastsQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Candida magnoliae
US-10-639-159-2

Query Match          26.7%; Score 378; DB 16; Length 283;
Best Local Similarity 36.8%; Pred. No. 1,6e-28;
Matches 99; Conservative 48; Mismatches 100; Indels 22; Gaps 8;

QY      8 FRPDGHLTVTVGACGGALAEALIKGLLAYGSDIALIDDOEKTAAKQAEYHKYATEELKXK 67
DB      29 FLKNGKVASITGSSSGIGYALAEAPQAGDAVALIWNSHDATGKAELATAKYGIV----- 82
QY      68 EYPRKGSYACDISDSDYTHKVFPAQVAKDPKPLHLVNTAG--YCENFPCEDYPAKNAEK 125
DB      83 ---KPKAKVANSSSDAVKQTIEQIKPFGHLDIVAN-AGIPWTKGAYIDODDDKHPDQ 138
QY      126 MYKVVLLLSLYS----QAFAPKLKEGIKGASVVLIGSMGSAIVNDPONOVYNNMSKXG 181
DB      139 VVDVDLKGVGIVAKHAGHFRFRFEKKEKKA-LVFTASMSGHIVNVPQFQATYNAAKAG 197
QY      182 VYHLAKTLACEWAKXNIRVNSINPGYIVGLTKNVINGNEELVNMISGIPQQRMSPEKE 241
DB      198 VRHFAKSLAVEPAP-ARVNSVSPGYINTIEDSFV---PQETQRKMSVLVPLGRGGETAE 253
QY      242 YTGAVLYLISESAAYTTGASILLVNGGFT 270
DB      254 LVGAYLFLASD-AGSYATGTDTIIVDGGYT 281

```

```

Db      139 VDDVDLKGVGIVAKHAGRHRREREKGGKGA-LVFPAISMGSCHIVNVPQQAATYNAKAG 197
QY      182 VIHLAKTLACWAKYKINIRVNSLNGYIYGPLTKNVINGNEELYNRYISGIPQORMSEPK 241
Db      198 VRHFAKSLAVEFAF-ARVNSVSPGYINTEISDFV---PQTONKMWSLVPLGRGETAE 253
QY      242 YTGAVLYLISSEASAYTTGASLLVDGGFT 270
Db      254 LVGAYVLFPLASD-AGSYATGTDTIIVDGGYT 281

RESULT 14
US-10-639-159-6
; Sequence 6, Application US/10639159
; Publication No. US20040137585A1
; GENERAL INFORMATION:
; APPLICANT: Davis, S. Christopher
; APPLICANT: Gray, John H.
; APPLICANT: Gruber, John M.
; APPLICANT: Hulsman, Gjaft W.
; APPLICANT: Ma, Steven K.
; APPLICANT: Newman, Lisa M.
; APPLICANT: Sheldon, Roger
; APPLICANT: Wang, Li A
; TITLE OF INVENTION: ENZYMATIC PROCESSES FOR THE PRODUCTION
; TITLE OF INVENTION: OF 4-SUBSTITUTED-3-HYDROXYBUTYRIC ACID DERIVATIVES
; FILE REFERENCE: 0339.210US
; CURRENT APPLICATION NUMBER: US/10/639,159
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: US 60/402,436
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 283
; TYPE: prt
; ORGANISM: Candida magnoliae
US-10-639-159-6

Query Match      26.7%; Score 378; DB 16; Length 283;
Best Local Similarity 36.8%; Pred. No.1.6e-28;
Matches 99; Conservative 48; Mismatches 100; Indels 22; Gaps 8;

QY      8 FFFDGLTIVTGACGGALBALIKGLAYGSDIALIDIOEKTAQKQAEYHKYATBELK 67
Db      29 FPLNGKVASITSSSGIGYALAEFAOVGADVAIWNSHDITGKAEALAKKYGV----- 82
QY      68 EYFPAKGSYACDISDDDTYHKVPAOVAKDPGKLPFLTVNTAG--YCFNFPCEPDYPAKNAEK 125
Db      83 ---KRAYKANVSSSDAYKQTEQIKPFGHLIDIVAN-AGIPLTKGAYIDODDDKHFDQ 138
QY      126 MYKNVLISGLYS----QAFAPLIIKEGIKGASVVLIGSMGSAIVNDPQNVVNMKAG 181
Db      139 VDDVDLKGVGIVAKHAGRHRREREKGGKGA-LVFPAISMGSCHIVNVPQQAATYNAKAG 197
QY      182 VIHLAKTLACWAKYKINIRVNSLNGYIYGPLTKNVINGNEELYNRYISGIPQORMSEPK 241
Db      198 VRHFAKSLAVEFAF-ARVNSVSPGYINTEISDFV---PQTONKMWSLVPLGRGETAE 253
QY      242 YTGAVLYLISSEASAYTTGASLLVDGGFT 270
Db      254 LVGAYVLFPLASD-AGSYATGTDTIIVDGGYT 281

RESULT 15
US-10-639-159-8
; Sequence 8, Application US/10639159
; Publication No. US20040137585A1
; GENERAL INFORMATION:
; APPLICANT: Davis, S. Christopher
; APPLICANT: Grate, John H.
; APPLICANT: Gray, David R.

```



```

; APPLICANT: Gruber, John M.
; APPLICANT: Hulsmann, Gfalt W.
; APPLICANT: Ma, Steven K.
; APPLICANT: Newman, Lisa M.
; APPLICANT: Sheldon, Roger
; APPLICANT: Wang, Li A
; TITLE OF INVENTION: OF 4-SUBSTITUTED-3-HYDROXYBUTYRIC ACID DERIVATIVES
; TITLE OF INVENTION: OF 4-SUBSTITUTED-3-HYDROXYBUTYRIC ACID DERIVATIVES
; PILE REFERENCE: 0339.210US
; CURRENT APPLICATION NUMBER: US/I0/639,159
; PRIOR FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: US 60/402,436
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 263
; TYPE: PRF
; ORGANISM: Candida magnoliae
; US-10-639-159-8

```

Query March	26.7%	Score 378;	DB 16;	Length 283;
Best Local Similarity	36.8%;	Pred. No. 1.6e-28;		
Matches 99;	Conservative 48;	Mismatches 100;	Indels 22;	Gaps 8

```

OY 8FFPDGHLITVTAQCGGLAEHLKIGLTAQSDMLDLIDQKTAQAQAEHKKYATEEHLK 67
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 29 FTLNGVVAISITSSSSGIGYALAEAPQVQGVDAIWNNSHATGAEHLAKKTYG----- 82
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 68 EYFKKGSYACIDISDPTVHKVPAQVADPGLPLHLVNTAG--YCENPCEEDYPAKNAEK 122
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 83 ---KKYAKYANVSSSDAVKQOTIEQQKDFGHLIDIVAN--AGIPWKCAIYIDQDDDKFDD 138
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 126 MYKNVLIGSLYXS---QAPAKPLIEGIGKASVYLIGSMGAILVNDPQNVVYNNMSKAG 181
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 139 VVDVVLKGVYAKAGRHFRFEFEKGGKGA-LVTPAASSGHIVANPQOATYNNAKAG 197
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 182 VIHLAKTLACWAKTINIRVNSLNPGYIYGLTNRQVINGNEELYNNRMTISGIPQOUMSEPK 241
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 198 VSHFPAKSLAVEAPAF-ARVNSVSPGYINTEISDFV---PQTQNKWMSLVPLGRGSETAE 255
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 242 YIGAVLYLISESAASYTNGASLLVNDGFT 270
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 254 LVGAYFLPLASD--AGSYATGTDLIDVGGYT 281
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

Search completed: July 9, 2005, 10:44:43
Job time : 163 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 10:20:26 ; Search time 41 Seconds

(without alignments)
638.316 Million cell updates/sec

Title: US-10-720-018-2

Perfect score: 1415

Sequence: 1 MTDTYTPFRFGHLLTVTGA.....SASVYTGASLLVDGGFTSW 272

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	798.5	56.4	278	2	D-arabinitol 2-deh
2	797	56.3	282	2	D-arabinitol 2-deh
3	402.5	28.4	257	2	oxidoreductase, sh
4	378	26.7	283	2	cardonyl reductase
5	376	26.6	262	2	short chain dehydr
6	376	26.6	262	2	oxidoreductase, sh
7	361	25.5	255	2	sorbitol utilizati
8	358.5	25.3	281	2	2-deoxy-D-gluconat
9	349	24.7	271	2	probable short cha
10	344	24.3	255	2	hypothetical prote
11	336	23.7	255	2	oxidoreductase, sh
12	332.5	22.8	255	2	probable short-chain
13	321.5	22.0	262	2	xylicol dehydrogen
14	312	22.0	255	2	hypothetical prote
15	305.5	21.6	267	2	hypothetical prote
16	305	21.6	261	2	short-chain dehydr
17	303.5	21.4	271	2	probable tropinone
18	302	21.3	254	2	dehydrogenase/redu
19	302	21.3	254	2	dehydrogenase/redu
20	301	21.3	268	2	probable tropinone
21	298	21.1	251	2	gluconate 5-dehydr
22	291	20.6	263	2	gluconate 5-dehydr
23	289	20.4	256	2	gluconate 5-dehydr
24	289	20.4	521	1	hypothetical prote
25	288.5	20.4	260	2	gluconate 5-dehydr
26	287.5	20.3	262	2	gluconate 5-dehydr
27	287.5	20.3	268	2	probable tropinone
28	287	20.3	298	2	gluconate 5-dehydr
29	287	20.3	298	2	gluconate 5-dehydr

30	285.5	20.2	262	2	C64711	probable tropinone
31	285	20.1	256	2	A10406	probable dehydroge
32	284	20.1	255	2	G82644	2,5-dichloro-2,5-c
33	281.5	19.9	262	2	G84694	probable tropinone
34	281.5	19.9	322	2	G84694	probable tropinone
35	281	19.9	246	2	H72219	3-oxocetyl-(acyl ca
36	281	19.9	268	2	G96016	probable gluconate
37	279	19.7	285	1	A65017	probable tropinone
38	278	19.6	269	2	C84693	probable tropinone
39	277	19.6	268	2	B84693	probable tropinone
40	277	19.6	285	2	A91041	probable oxidoredu
41	277	19.6	285	2	D85885	probable oxidoredu
42	276.5	19.5	253	2	C91091	2-deoxy-D-gluconat
43	275.5	19.5	253	2	F85936	2-deoxy-D-gluconat
44	275	19.4	267	2	G97220	short-chain alcohoh
45	274.5	19.4	260	2	A84695	probable tropinone

ALIGNMENTS

RESULT 1

S57351 D-arabinitol 2-dehydrogenase (EC 1.1.1.-) - Yeast (Pichia stipitis)

C:Species: Pichia stipitis

C:Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004

C:Accession: S57351; S57352

R:Hallborn, J.; Walfriedson, M.; Penttilae, M.; Keranen, S.; Hann-Haegderl, B.

A>Title: A short-chain dehydrogenase gene from Pichia stipitis having D-arabinitol dehydr

A:Reference number: S57351; MUID:9609013; PMID:7483848

A:Accession: S57351

A:Molecule type: DNA

A:Residues: 1-278 <HML>

A:Cross-references: UNIPROT:P50167; EMBL:Z46866; NID:G758241; PIDN:CAA86939.1; PID:G7631

A:Accession: S57352

A:Molecule type: mRNA

A:Residues: 1-278 <HAM>

A:Cross-references: EMBL:Z46866; NID:G758241; PIDN:CAA86939.1; PID:G763164

A:Experimental source: strain CBS 6054

C:Genetics:

A:Gene: ARDH

C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

C:Keywords: NAD; oxidoreductase

F:17-212/Domain: short-chain alcohol dehydrogenase homology <SADH>

F:17-47/Region: beta-alpha-beta NAD nucleotide-binding fold

Query Match 56.4%; Score 798.5; DB 2; Length 278;
Best Local Similarity 54.3%; Pred. No. 1.1e-55;
Matches 151; Conservative 54; Mismatches 64; Indels 9; Gaps 3;

QY	3	DY-----IPFRFGHLLTVTGA	CGGLAEALIKLLAVGSLALDDDEKTAQKQAEYH	57
DB	2	DYSYANVVPNFRLLGRLLIITGSGGLAAVLSRLLAQADVALIDNLETKKAANEVL	61	
QY	58	KVATESELKKEVPMG---SYACDISDPYHAKVPAOVAOFGKLPHLVVTAGYCFENP	114	
DB	62	GWGSEYTLKGHSAIKGVASMSCHIGDAEAVDAFFSSINEHGKTAADLLITATGTCENFP	121	
QY	115	CEYPAKNAEKMYKNLLGLSYQAFAKPLIKGIGASVLLIGSMGAILVNDPQNV	174	
DB	122	AETYPATNAESIMKVMGLGSPYVQSFPARPLIQNNLNG-SIILIGSMGTVNDPQCM	180	
QY	175	YMSKAGVIHLAKTLAEAKYNNRNSLMPGYLYGSLTKNVINGNBEYRWJSGIFPOQ	234	
DB	181	YMSKAGVIHLVRSLACEMAKYNNRNTLSPGYILPLTRNVISGHTMEKAWESKLPK	240	
QY	235	RMSPEKEVIGAVLYLSESAAVYTTGASLLVDGGFTSW	272	
DB	241	RMAPKPEFVSILYLASETASSVYTTGHNLVVDGGYECW	278	

RESULT 2

RESULT 5
AG2809
Short chain dehydrogenase Atu1897 [Imported] - Agrobacterium tumefaciens (strain C58, Du
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AG2809
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kucyavin, T.; Levy, R.; Li, M.; Mclell
; Karp, P.; Romero, P.; Zhang, S.
Science 294 2317-2323, 2001
A:Authors: Yoo, H.; Tso, Y.; Biddle, P.; Jung, M.; Krepsan, W.; Perry, M.; Gordon-Kamm,
ser, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AG2809
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-262 <KUR>
A:Cross-references: UNIPROT:Q8UB64; GB:AE008688; PIDN:AAL42893.1; PID:g17740346; GSPDB:C
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu1897
A:Map position: circular chromosome
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 26.6%; Score 376; DB 2; Length 262;
Best Local Similarity 35.8%; Pred. No. 2.5e-22;
Matches 96; Conservative 43; Mismatches 109; Indels 20; Gaps 7;

QY 8 FRPDGHLTIYVAGCGGLAEALIKGLAYGSDIALDIDOEKTAQAQAYHKYATEBLTK 67
DB 12 FDLGQVALVYAGSGGIGORIAMGLAQSAGAVALLDRRTDGLAQTAQFIARAGR----- 66
QY 68 EVPKMGSAACISDSDTYHKVPAQVAKDFGLPLHLVNTAGYCEHPCEDEYPAKAEKV 127
DB 67 ---KSIQIAADVTSKQALTDVAARQAELGALSL-AVNAAGIANNPAREMBESQFQTM 122
QY 128 KVNLLGSLVYSAQAPKPLIKEGKIGASVVLIGSMGALVNDPQNVVYNSKAGVIHLAK 187
DB 123 DINLKGVLSCQAEANMLKNG-RGA-IVNIAISMGIIVNRGLMCHYNASKAGVIHNSK 180
QY 188 TLACEMAKYINRVNSLNGYIYGLPTKVNINGNEELVNR---WISGIPQOMSEPKXYIG 244
DB 181 SMAMWVGRIKRVNTISPGYATPW-----NTRPEMVHQTFLFEQOTMOMAGVDENVG 235
QY 245 AVLYLLSESAASYTTGASLLVDGFTSW 272
DB 236 PAIFLLSD-AASFVTVGVDLLVDGFCFW 262

RESULT 6
B97588
oxidoreductase, short chain dehydrogenase/reductase family [Imported] - Agrobacterium tu
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: B97588
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Mollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: B97588
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-262 <KUR>
A:Cross-references: UNIPROT:Q8UB64; GB:AE007869; PIDN:AAK87659.1; PID:g15157014; GSPDB:C
C:Genetics:
A:Gene: AGR C3482
A:Map position: circular chromosome
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 26.6%; Score 376; DB 2; Length 262;
Best Local Similarity 35.8%; Pred. No. 2.5e-22;
Matches 96; Conservative 43; Mismatches 109; Indels 20; Gaps 7;

QY 8 FRPDGHLTIYVAGCGGLAEALIKGLAYGSDIALDIDOEKTAQAQAYHKYATEBLTK 67
DB 12 FDLGQVALVYAGSGGIGORIAMGLAQSAGAVALLDRRTDGLAQTAQFIARAGR----- 66
QY 68 EVPKMGSAACISDSDTYHKVPAQVAKDFGLPLHLVNTAGYCEHPCEDEYPAKAEKV 127
DB 67 ---KSIQIAADVTSKQALTDVAARQAELGALSL-AVNAAGIANNPAREMBESQFQTM 122
QY 128 KVNLLGSLVYSAQAPKPLIKEGKIGASVVLIGSMGALVNDPQNVVYNSKAGVIHLAK 187
DB 123 DINLKGVLSCQAEANMLKNG-RGA-IVNIAISMGIIVNRGLMCHYNASKAGVIHNSK 180
QY 188 TLACEMAKYINRVNSLNGYIYGLPTKVNINGNEELVNR---WISGIPQOMSEPKXYIG 244
DB 181 SMAMWVGRIKRVNTISPGYATPW-----NTRPEMVHQTFLFEQOTMOMAGVDENVG 235
QY 245 AVLYLLSESAASYTTGASLLVDGFTSW 272
DB 236 PAIFLLSD-AASFVTVGVDLLVDGFCFW 262

RESULT 7
T39164
sorbitol utilization protein soul - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T39164
R:McLean, J.; Harris, D.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, February 1998
A:Reference number: Z21831
A:Accession: T39164
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-255 <MCU>
A:Cross-references: UNIPROT:Q9Y629; EMBL:AL021817; PIDN:CAB40197.1; GSPDB:GN00066; SPDB:
A:Experimental source: strain 972h-, cosmid c8E11
C:Genetics:
A:Gene: SPAC8E11.10
A:Map position: 1
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 25.5%; Score 361; DB 2; Length 255;
Best Local Similarity 36.6%; Pred. No. 3.7e-21;
Matches 98; Conservative 45; Mismatches 103; Indels 22; Gaps 10;

QY 8 FRPDGHLTIYVAGCGGLAEALIKGLAYGSDIALDIDOEKTAQAQAYHKYATEBLTK 67
DB 5 FSLKGTLLITGSGGIGFSTAKAPAAAGSNVGLL-YGRNKKALFYA-----AELRDK 56
QY 68 EVPKMGSAACISDSDTYHKVPAQVAKDF-GKPLHLVNTAGYCEHPCEDEYPAKAEKV 124
DB 57 HGVQAKAYSCPIERSAVIETNQAVEBGRDLVMIANGIALPHLSLED---KNEDIW 113
QY 125 -KMKVNVLLGSLVYSAQAPKPLIKEGKIGASVVLIGSMGALVNDPQNVVYNSKAGVI 183
DB 114 TKVGVINLNGYVYPAQAAGHFFKQKQ-KG-SLIETASMSGITANWPOOMASYHATKAVK 171
QY 184 HLAATLCEMAKVIIRVNSLNGYIYGLPTKVNINGNEELVNRWISGIPQOMSEPKXYI 243
DB 172 HLAALAVENAPF-ARNSVSPGIITDITL---LYADENLRKKKKEYTPQARIGLPDEL 227
QY 244 GAVLYLLSESAASYTTGASLLVDGFTSW 271
DB 228 GAVLYLLSD-AASYCTGSDIIVDGYS 254

RESULT 8
F69400
2-deoxy-D-gluconate 3-dehydrogenase (kduD) homolog - Archaeoglobus fulgidus

C/Species: Archaeoglobus fulgidus
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C/Accession: F69400
R/Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A/Authors: Ueberbach, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A/Reference number: A69250; MUID:98049343; PMID:9389475
A/Accession: F69400
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-281 <KLE>
A/Cross-references: UNIPROT:O29061; GB:A601021; GB:A600782; NID:g2689344; PIDN:AAB9004
C/Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology
F/36-216/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match 25.3%; Score 358.5; DB 2; Length 281;
Best Local Similarity 33.3%; Pred. No. 6.6e-21;
Matches 91; Conservative 53; Mismatches 106; Indels 23; Gaps 9;

Qy 4 YIPFRDGLTIVTGACGGLAELIKGLAYGSDIALDIDOKTKAKOAEYHKYATEE 63
Db 27 YRKMFDLTKVAIVTATGATGCGPGLADPGCDVVVV-----GRRL-----VLEK 74
Qy 64 LKKEVPMGSA-----CDISDSTVHKVPQAVKDFGLHLVNTGCEMPCDEY 119
Db 75 LK-SIEELGGRALAVKCDITSEEDVANLVKRYVEEFERIDI-LVNCGINIPKAEY 132
Qy 120 AKNAEKVKNVLGSLVYSQAFKPLKEGKASVVLIGMSGAIYVDPQNVYNNMSK 179
Db 133 LEDNKKVDAVTVGVLYCREVGKRMVQ--NGGKINIVSSVRS--YEMPNYLAAGCSK 189
Qy 180 AGVHLAKTACEMAKNIRVNSLNGYIYGLTKNVINGBELYNNISGIPQORSEP 239
Db 190 AAVNMITQGLACEMAKNIVLVAIPVIAPIPLTHIMK-DPELSKTKWSRILLGRMGYP 248
Qy 240 KEYGAVVYLSESAASYTTGASLLVDDGFTSW 272
Db 249 DDLIGAVVFASD-ASNFTVGLIYDGGVTSW 280

RESULT 9

AC0157
Probably short chain oxidoreductase YPO1287 [imported] - Yersinia pestis (strain CO92)
C/Species: Yersinia pestis
C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C/Accession: AC0157
R/Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tilball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett,
Nature 413, 523-527, 2001
A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A/Reference number: AB0001; MUID:21470413; PMID:11586360
A/Accession: AC0157
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-271 <KUR>
A/Cross-references: UNIPROT:O8ZGK4; GB:AL590842; PIDN:CAC90118.1; PID:g15979338; GSPDB:C
A/Genetics:
C/Genetics:
C/Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 24.7%; Score 349; DB 2; Length 271;
Best Local Similarity 34.3%; Pred. No. 3.5e-20;
Matches 91; Conservative 44; Mismatches 102; Indels 28; Gaps 7;

Qy 12 GHLTVTGACGGLAELIKGLAYGSDIALDIDOKTKAK-----QAEYHKYATEELK 67
Db 29 GRLAVTGGGCGGAGALAGLAATGAEVICTDISERRAAATQALNAKGVNRAEGL--- 85

Qy 68 EYRPMGSIYACDISDSTVHKVPQAVKDFGLPLH-LVNTAGCENPCDEYPAKNAEK 126
Db 86 -----DYVDSALIDALNAALP-----PLDVLVCNAGLVTHTPAEEMTDADWDKV 129
Qy 127 VKNVLGSLVSGAFAPKPLIKEGIKGASVYLIGMSGAIYVDPQNVYNNMSKAGVYHLA 186
Db 130 IAVNLGVFRTGCGFGRRLKEAG-RGSIINISISQIVNVQPOCHNPAASGAGVHLL 187
Qy 187 KTLACEMAKNIRVNSLNGYIYGLTKNVINGBELYNNISGIPQORSEPKEKYI 246
Db 188 KSLAVEMATGAVVNAVAFYIETPLQG-LTSQPGVSRHMDMTBMGRLSPEHIAV 246
Qy 247 LYLSESAASYTTGASLLVDDGFTS 271
Db 247 QPLASE-ASSLVLGSIITADAGYTS 270

RESULT 10

D70635
hypothetical protein Rv1928c - Mycobacterium tuberculosis (strain H37RV)
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C/Accession: D70635
R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Fieldwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A/Authors: Squires, R.; Suleston, J.E.; Taylor, K.; Whitehead, S.; Barrett, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; MUID:9825987; PMID:9634220
A/Accession: D70635
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-255 <COL>
A/Cross-references: UNIPROT:P95286; GB:Z84498; GB:AL123456; NID:g3261701; PIDN:CAB06498.1
A/Experimental source: strain H37RV
C/Genetics:
A/Genetics:
C/Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology
F/12-194/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match 24.3%; Score 344; DB 2; Length 255;
Best Local Similarity 33.7%; Pred. No. 8.1e-20;
Matches 90; Conservative 45; Mismatches 108; Indels 24; Gaps 8;

Qy 8 FRPGLTIVTGACGGLAELIKGLAYGSDIALDIDOKTKAKOAEYHKYATEELK 67
Db 7 FDLGKALALITGASTGIGKVALAYVERAGVAL-----AARHLDLEKL-AD 53
Qy 68 EVPMKGS-----YACDISDSTVHKVPQAVKDFGLPLHVLNTGCEMPCDEYPAKNA 123
Db 54 EIGTSGKVPVCCDVSOHQOVTSMLDQVTAELGDIIVACN-AGIITVTMLMPEEF 112
Qy 124 EKMTKNVLGSLVYSQAFKPLIKEGKASVVLIGMSGAIYVDPQNVYNNMSKAGV 183
Db 113 QRLONTVNTGVLFLQAALAKMKVQK-QGQVITNTASGSHIIVPQOVSHYKSKAAV 171
Qy 184 HLAKTACEMAKNIRVNSLNGYIYGLTKNVINGBELYNNISGIPQORSEPKEKY 243
Db 172 HLTKAMAVELAPKIRVNSVPGYI---LTB-LVEPTVEYPLMEPKPLGLRLRPBELA 227
Qy 244 GAVVYLSESAASYTTGASLLVDDGFT 270
Db 228 GVIYVLAEE-ASSVMTGSDIYIDGGYT 253

RESULT 11

D72377
oxidoreductase, short chain dehydrogenase/reductase family - Thermotoga maritima (strain
C/Species: Thermotoga maritima
C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C/Accession: D72377
R/Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey

Garratt, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.,
C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:99287316; PMID:10360571
A:Accession: D72377
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-255 <ARN>
A:Cross-references: UNIPROT:Q9WYS2; GB:AE001722; GB:AE000512; NID:94980938; PIDN:AAD3552
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM0441
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 23.7%; Score 336; DB 2; Length 255;
Best Local Similarity 33.5%; Pred. No. 3.5e-19;
Matches 89; Conservative 49; Mismatches 110; Indels 18; Gaps 9;

Qy 8 PRPGHLLTVTGACCGALBALIKGLAYGSDIALDIDQEKTAQAQAEYHKYATELKL 66
Db 5 FDLGRVALVTYGGSRGLGFGIAOGIAEAGCSVVVASRLBEASBAQKLTKEYGVETM-- 62

Qy 67 KEVPMGYSACDISDPTVHVFAQVAKDFGLPLHLVNTAGYCENFPCEDYPAKNAEM 126
Db 63 -----AFRCDSVSYEEVKKLLEAVKEKFGKLDLTVNNAAGINRRHAEPEPLDEFRQV 114

Qy 127 KVNLLGSLYV-SQAFAPLKEGIGKASVVLIGMSGAIYNDPONQVYNNKAGVTHL 185
Db 115 IEVNLFGTYVYCREAFS--LRES-DNPSINIGSLTYEETMP-NISAYASKGVASL 170

Qy 186 AKTLACWAKYNIRVNSLPGYITGLTKVNINGNELYNMISGIPQORMSEPKRYGA 245
Db 171 TKALAKWEGRGYRIVNVIAFGMYRTKMTAEVFSDEKL-DYMLKRIPLGRGVPEDLKGV 229

Qy 246 VLYLSESAASYTGASLLVDDGFTS 271
Db 230 AVFLASEE-AKTVGQILFVDGQMTA 254

RESULT 12
D83416
probable short-chain dehydrogenase PA1828 (imported) - Pseudomonas aeruginosa (strain PA
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: D83416
R:Stover, C.K.; Pham, X.Q.; Eryin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Laidig, K.; Lim,
., Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: D83416
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-255 <STO>
A:Cross-references: UNIPROT:Q912R7; GB:AE004609; GB:AE004091; NID:99947810; PIDN:AAG0521
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA1828
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 22.8%; Score 322.5; DB 2; Length 255;
Best Local Similarity 34.7%; Pred. No. 4.1e-16;
Matches 93; Conservative 48; Mismatches 104; Indels 23; Gaps 10;

Qy 8 PRPGHLLTVTGACCGALBALIKGLAYGSDIALDIDQEKTAQAQAEYHKYATELKL 67
Db 7 FDLGRKAPVSGARGIGBALAKLAQGAHVI--VSSRKIDCCQ-----AVADAITA 57

Qy 68 EVPMGYSACDISDPTVHVFAQVAKDFGLPLHLVNTAG-----XENFPCEDYPAKNA 123
Db 58 EGGKATAIACHTGEMEQIONVFAIRDFGRLLD-LVNNNAATNPQFC-NVLETLGA--F 113

Qy 124 ERMVKNLLGSLYVSOAFAPLKEGIGKASVVLIGMSGAIYNDPONQVYNNKAGVY 183
Db 114 OKTVDNIRGVYFMSIEGKLMKEHG--GGSILINVASING--VSPGEFGIYSVTKAIV 169

Qy 184 HLAKLACWAKYNIRVNSLPGYITGLTKVNINGNELYNMISGIPQORMSEPKRY 243
Db 170 SMTKVFACQAFQIRCNALLPGLDTYFASALVK-NDAIRNLQRIPLKRVAPSEWA 228

Qy 244 GAVLYLSESAASYTGASLLVDDGFTS 271
Db 229 GAVLYLASEE-ASSTTVGVALVDDGFTS 255

RESULT 13
JC7939
xyllitol dehydrogenase (BC 1.1.1.-) - Gluconobacter oxydans (Strain ATCC621)
C:Species: Gluconobacter oxydans (Strain ATCC621)
C:Date: 22-Jun-2003 #sequence_revision 22-Jun-2003 #text_change 07-Jul-2003
C:Accession: JC7939; PC7223
R:Sugiyama, M.; Suzuki, S.; Tomouchi, N.; Yokozaki, K.
Biosci. Biotechnol. Biochem. 67, 584-591, 2003
A:Title: Cloning of the xyllitol dehydrogenase gene from Gluconobacter oxydans and improv
A:Reference number: JC7939; MUID:22608835; PMID:12723607
A:Accession: JC7939
A:Molecule type: DNA
A:Residues: 1-262 <SUG>
A:Cross-references: DDBJ:AB091690
A:Accession: PC7223
A:Molecule type: protein
A:Residues: 2-26 <SUG>
A:Comment: This enzyme is an NADH-dependent enzyme and belongs to the short-chain dehydr
xyllitol from D-xylose or D-arabitol.
C:Genetics:
A:Gene: xdh
C:Keywords: NADH-dependent enzyme; short-chain dehydrogenase/reductase family; xyllitol d

Query Match 22.7%; Score 321.5; DB 2; Length 262;
Best Local Similarity 31.0%; Pred. No. 5e-18;
Matches 86; Conservative 53; Mismatches 103; Indels 35; Gaps 10;

Qy 9 PRPGHLLTVTGACCGALBALIKGLAYGSDIALDIDQEKTAQAQAEYHKYATELKL 68
Db 4 KFNKVCVLVAGAGNIGLALTRLAEEGTALIDMNRREALKFAASVREKVE----- 57

Qy 69 VPKMGYSACDISDPTVHVFAQVAKDFGLPLHLVNTAGCENF-PCEDYPAKNAEM 127
Db 58 ---ARSTVCVVTSEAVITGVDSVVRDFKIDF-LFNNAAGYQAFAPVODPSPDDFARVL 113

Qy 128 KVNLLGSLYVSOAFAPLKEGIGKASVVLIGMSGAIYNDPONQVYNNKAGVTHLAK 187
Db 114 TIVNTAGFHLKAVSRQMITQNY--GRIVTASNAG--VGGPRMAAYGTSKAIILTE 169

Qy 188 TLACWAKYNIRVNSLPGYITGP-----LTKVNINGNE-----ELYNMISGI 231
Db 170 TAAIDLALPYNIRVNAISPGYV-CPGFMWERQVELQAKV--GSQYFSTDPKVVAQOMIGSV 226

Qy 232 PQORMSEPKRYGAVLYLSESAASYTGASLLVDDG 268
Db 227 PMRRYGDINEIPGVVAFLLGDD-SSFMTGVNLPLAGG 262

RESULT 14
T15987
hypothetical protein F09E10.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T15987
R:Geiselt, C.; Gattung, S.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans csmid F09E10.
A:Reference number: Z18443
A:Accession: T15987

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-255 <Gen>
A:Cross-references: UNIPROT:Q19246; EMBL:U41749; NID:g1118144; PID:g1118145; PIDN:AA524
A:Experimental source: strain Bristol N2; clone F09E10
C:Genetics:
A:Gene: CESP:F09E10.3
A:Map position: X
A:Insertions: 5/2; 101/3; 126/3; 184/3
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 22.0%; Score 312; DB 2; Length 255;
Best Local Similarity 33.2%; Pred. No. 2.7e-17;
Matches 85; Conservative 47; Mismatches 100; Indels 24; Gaps 7;

Qy 18 TGACGGALAEALIKGLAAGSDIALDIDDEKTAQAQAEYHKVATEBELKKEVP---KMG 74
Db 20 SGASGIGKAIQTLAKGARVAVADLDGNA---TAKALPASQSHS 66

Qy 75 YACDISDVTAKVPAQVAKDFGKLEHLVNTAGYCENPCEDYPAKNAERKVKVNLGS 134
Db 67 PACDVSNDV-KGLSHVKSIG--TPSLVNCAGITKSTLLKMQEWDVVKVNLGV 124

Qy 135 LVYSQAFKAPLIKEGIKGASVVLIGMSGAIVNDPQNOVYTNMSKAGVIHLAKTLACWA 194
Db 125 FHVSOAFKASVDNNHPLSLIINSSIVGKMGNFQTN--YAATKAGVIGFTKSAKELA 182

Qy 195 KYNRVNSLNPXYIGPLTKVINGNELYNRMISGIPQRMSEPEKVIYGAIVLYLSESA 254
Db 183 KKNRVNNAVLPGFKTPTEAM---PTVLALICKGIPMRGEMETANSVLYLSD-L 238

Qy 255 ASYTGASILVDGGFT 270
Db 239 SSYTGATLEVTGFS 254

RESULT 15

hypothetical protein alr4456 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A>Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AH2362
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuchih
Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A>Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AH2362
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-267 <Gen>
A:Cross-references: UNIPROT:Q8YNV5; GB:BA000019; PIDN:BA676155.1; PID:g17133592; GSPDB:C
A:Experimental source: strain PCC 7120
A:Gene: alr4456
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 21.6%; Score 305.5; DB 2; Length 267;
Best Local Similarity 30.7%; Pred. No. 9.5e-17;
Matches 80; Conservative 55; Mismatches 109; Indels 17; Gaps 7;

Qy 12 GHLTVTGACGALAEALIKGLAAGSDIALDIDDEKTAQAQAEYHKVATEBELKKEVPK 71
Db 22 GRKALITGKTGIGLAIQAEFLALGAEVIVARNARAEIQOMKAWH-----SAGK 71

Qy 72 MGSYACDISDVTAKVPAQVAKDFGKLEHLVNTAGYCENPCEDYPAKNAERKVKVNL 131
Db 72 VHGAVADVSTSEGRQMDLVYVSKTFGEIPI-LVNNVGTNIRKKATDYTEEFPAIFQINL 130

Qy 132 LQSLVSOAFKAPLIKEGIKGASVVLIGMSGAIVNDPQNOVYTNMSKAGVIHLAKTLAC 191
Db 131 TSIFELSRIF-YPLLKTS-KNSSIVIGSVAGLI--SVRTGAPYGTAKAALVOLTRSLAV 186

Qy 192 EMAKYNIRVNSLNPXYIGPLTKVINGNELYNRMISGIPQRMSEPEKVIYGAIVLYLS 251
Db 187 EMADDGIRVNAIAPWFIQTLTEPLIN-NPETLSAVLSRTPMKRVGQPEE-VASLTAFLC 244

Qy 252 ESAASYTTGASLVVDGGFTSW 272
Db 245 MPTASYTTGQCIADVDDGFIAP 265

Search completed: July 9, 2005, 10:31:43
Job time : 42 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 10:19:36 ; Search time 178 Seconds

(without alignments)
782.503 Million cell updates/sec

Title: US-10-720-018-2

Perfect score: 1415

Sequence: 1 MTDIYPTFRFDGHLITVGA.....SAASYTGASLLVDGFTSW 272

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1415	100.0	272	Q70FD1	Q70FD1 ambrosiozym
2	800	56.5	281	ARDH_CANAL	P43066 candida alb
3	798.5	56.4	278	1 ARDH_PICST	P50167 pichia stip
4	797	56.3	282	1 ARDH_CANTR	P50166 candida tro
5	769.5	54.4	278	2 O6BY61	O6BY61 debaryomyce
6	659.5	47.3	314	2 O6CX46	O6CX46 kluyveromyce
7	528.5	37.3	343	2 O6C367	O6C367 yarrowia li
8	510.5	36.1	368	2 O6T5L8	O6T5L8 emericella
9	501	35.4	371	2 O7SE47	O7SE47 neurospora
10	465	32.9	291	2 O6C6W8	O6C6W8 yarrowia li
11	407	28.8	259	2 O6ESR1	O6ESR1 manheimia
12	402.5	28.4	257	2 O9WYD3	O9WYD3 thermotoga
13	399.5	28.2	291	2 O6BN16	O6BN16 debaryomyce
14	393	27.8	266	1 DCKR_TRIRE	O8nk50 trichoderma
15	392.5	27.7	257	2 O89QK5	O89QK5 bradyrhizob
16	389.5	27.5	364	2 O7RZD4	O7RZD4 neurospora
17	383.5	27.1	279	2 O6BQ25	O6BQ25 debaryomyce
18	382	27.0	281	1 SOU1_CANAL	P87218 candida alb
19	381	26.9	256	2 O92MP6	O92MP6 rhizobium m
20	378	26.7	283	2 O9C4B3	O9C4B3 candida mag
21	377.5	26.7	257	2 O92MR6	O92MR6 rhizobium m
22	376	26.6	262	2 O8UE64	O8UE64 agrobacteri
23	374	26.4	280	2 O6CM06	O6CM06 kluyveromyce
24	368.5	26.0	280	1 SOU2_CANAL	P87218 candida alb
25	361	25.5	255	2 O9Y629	O9Y629 schizosacch
26	360.5	25.5	265	2 O86ZP3	O86ZP3 gibberella
27	359	25.4	340	2 O7SDM3	O7SDM3 neurospora
28	358.5	25.3	266	2 O96W29	O96W29 cladosporiu
29	358	25.3	281	2 O29061	O29061 archaeglob
30	357	25.2	254	2 O8B3G6	O8B3G6 pseudomonas
31	356	25.2	271	2 O66CT3	O66CT3 yersinia ps

32	354.5	25.1	266	2	O8J231	O8J231 alternaria
33	350	24.7	271	2	O7S2T4	O7S2T4 neurospora
34	349	24.7	271	2	O8ZGK4	O8ZGK4 yersinia pe
35	348.5	24.6	255	2	O8COE7	O8COE7 staphylococ
36	348.5	24.6	257	2	O8FZ65	O8FZ65 bruceella su
37	347.5	24.6	259	2	O6BN65	O6BN65 debaryomyce
38	347	24.5	256	2	O92MR3	O92MR3 rhizobium m
39	345	24.4	273	2	O89H04	O89H04 bradyrhizob
40	344	24.3	255	2	P95286	P95286 mycobacteri
41	344	24.3	255	2	O7TZ97	O7TZ97 mycobacteri
42	343	24.2	268	2	O986J1	O986J1 rhizobium l
43	341.5	24.1	256	2	O73ZFA	O73ZFA mycobacteri
44	340.5	24.1	278	2	O6CE59	O6CE59 yarrowia li
45	339	24.0	285	2	O6BQ28	O6BQ28 debaryomyce

ALIGNMENTS

RESULT 1									
ID	Q70FD1	PRELIMINARY;	PRT;	272	AA.				
AC	Q70FD1								
DT	05-JUL-2004	(TRENBLrel. 27, Created)							
DT	05-JUL-2004	(TRENBLrel. 27, Last sequence update)							
DT	05-JUL-2004	(TRENBLrel. 27, Last annotation update)							
DE	NADH L-xyulose reductase (Fragment).								
GN	Name=alx1;								
OS	Ambrosiozyma monospora.								
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;								
OC	Saccharomycetales; Saccharomycopsidaceae; Ambrosiozyma.								
OX	NCBI_TaxID=43982;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN=NREL_Y-1484;								
FX	PubMed=14736891; DOI=10.1074/jbc.M312533200;								
RA	Verho R., Putkonen M., Londeborough J., Penttilae M., Richard P.;								
RT	"A novel NADH linked L-xyulose reductase in the L-arabinose catabolic								
RT	pathway of yeast.";								
RL	J. Biol. Chem. 279:14746-14751 (2004).								
DR	EMBL; AJ583159; CAE47547.1; -.								
DR	HSSP; P50162; IAE1.								
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.								
DR	GO; GO:0008152; P:metabolism; IEA.								
DR	InterPro; IPR002198; ADH_short.								
DR	InterPro; IPR002347; Adh_short_C2.								
DR	InterPro; IPR002110; ANK.								
DR	Pfam; PF00106; adh_short; 1.								
DR	PRINTS; PRO1415; ANKYRIN.								
DR	PRINTS; PRO0081; GDRDH.								
DR	PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.								
FT	NON TER 272								
SO	SEQUENCE 272 AA; 29636 MW; 3338DB95ECC473D CRC64;								
Query Match 100.0%; Score 1415; DB 2; Length 272;									
Best Local Similarity 100.0%; Pred. No. 7.3e-107;									
Matches 272; Conservative 0; Mismatches 0; Indels 0; Gaps 0									
Qy	1	MTDVIPTFRFDGHLITVGA	CGGLAEALIKGLAYGSDIALLDIDDKETAKQAQEHKRYA	60					
Db	1	MTDVIPTFRFDGHLITVGA	CGGLAEALIKGLAYGSDIALLDIDDKETAKQAQEHKRYA	60					
Qy	61	TEELKKEVPKMSVACDISDPTVHKVFAQVAKD	RGKLPPLHLVNTAGYCENPFCEYPA	120					
Db	61	TEELKKEVPKMSVACDISDPTVHKVFAQVAKD	RGKLPPLHLVNTAGYCENPFCEYPA	120					
Qy	121	KNAEKVKVNLGLSLVYSAQFAKPLIKEGIKGASVVL	IGSMGAIYNDPQNVVYNNSKA	180					
Db	121	KNAEKVKVNLGLSLVYSAQFAKPLIKEGIKGASVVL	IGSMGAIYNDPQNVVYNNSKA	180					
Qy	181	GVVHLAKTLACFAKAKNIRVNSLNPQYITGPLTKAN	YINGNEELYNRWISGIPQQRSEPK	240					
Db	181	GVVHLAKTLACFAKAKNIRVNSLNPQYITGPLTKAN	YINGNEELYNRWISGIPQQRSEPK	240					

QY 241 EYIGAVLYLSESAASYTTGASLIVDGGFTSM 272
 DB 241 EYIGAVLYLSESAASYTTGASLIVDGGFTSM 272

RESULT 2

ARDH CANAL ID ARDH CANAL STANDARD; PRT; 281 AA.
 AC P43066; 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE D-arabinitol 2-dehydrogenase [ribulose forming] (EC 1.1.1.250) (ARDH).
 GN Name=ARDH; Synonyms=ARDH;
 OS Candida albicans (Yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 CC NCBI_TaxID=5476;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MO-1;
 RX MEDLINE=94012494; PubMed=8407803;
 RA Wong B., Murray J.S., Castellanos M., Croen K.D.;
 RT "D-arabitol metabolism in Candida albicans: studies of the
 biosynthetic pathway and the gene that encodes NAD-dependent D-
 arabitol dehydrogenase.";
 RL J. Bacteriol. 175:6314-6320(1993).
 CC -1- CATALYTIC ACTIVITY: D-arabinitol + NAD(+) = D-ribulose + NADH.
 CC -1- PATHWAY: D-arabinitol biosynthesis; last step.
 CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
 (SDR) family.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; L16227; AAC37430.1; -.
 DR HSSP; Q9ZFY9; 1PK8.
 DR InterPro; IPR002198; ADH_short.
 DR InterPro; IPR002347; Adh_short_C2.
 DR Pfam; PF00106; adh_short_1.
 DR PRINTS; PR00081; GDRDH.
 DR PRINTS; PR00080; SDRPAMILY.
 DR TIGRFAMs; TIGR01830; 3oxo_ACP_reduc; 1.
 DR TIGRFAMs; TIGR01831; fabG_rel; 1.
 DR TIGRFAMs; TIGR01832; kduJ; 1.
 DR PROSITE; PS00061; ADH_SHORT; 1.
 KM NAD; Oxidoreductase.
 FT NP_BIND 25
 FT ACT_SITE 184 184 NAD (By similarity).
 FT BINDING 169 169 Proton acceptor (By similarity).
 FT BINDING 169 169 Substrate (By similarity).
 SQ SEQUENCE 281 AA; 30643 MW; B302A6411591F892 CRC64;

Query Match 56.5%; Score 800; DB 1; Length 281;
 Best Local Similarity 55.7%; Pred. No. 8e-57;
 Matches 151; Conservative 55; Mismatches 61; Indels 4; Gaps 2;

QY 5 IPTFRPDGHLTVTGACGGLAEALIKGLAYGSDIALIDDOEKTPAKQAQYHKAATEEL 64
 DB 12 VPSFRIDGKLVITIGSGGLAAVSRALAKGADVALLVDMNLEKRTQAARDVLQMGEEQM 71
 QY 65 KKK---EVRKMGSYACDISDSDTVHKVPAQVAKDFEKLPLHLVNTAGYGFCEPDYPAK 121
 DB 72 KKGYSERPIQGVAMSGNIDDAEAVDLTFALINEHHKISSVLTNTAGVAFENPAAEYPAK 131
 QY 122 NAEKMYKVALLSGLYVSOAPAKFLKEGKASVLLIGMSGALIVNDPQNVVYVYNSKAG 181
 DB 132 NENIMKVVGLGFSVYVSOAFARPLIONMTG--SIILIGSGTIVNDPQPOCMYNSKAG 190

QY 182 VIHAKTLACEMAKYNIIRVNSINPGYIGPLTKQVINGNEELYNRWISGIPQORMSEPYE 241
 DB 191 VIHAKSLACEMAKYNIIRVNTISPGYILPLFTRNVISGHTKMEWSKIPMKRMAEPYE 250
 QY 242 YIGAVLYLSESAASYTTGASLIVDGGFTSM 272
 DB 251 FVGSILYLAESASASYTTGHLVVDGGYEGCW 281

RESULT 3

ARDH PICST ID ARDH PICST STANDARD; PRT; 278 AA.
 AC P50167;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE D-arabinitol 2-dehydrogenase [ribulose forming] (EC 1.1.1.250) (ARDH).
 GN Name=ARDH;
 OS Pichia stipitis (Yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Pichia.
 CC NCBI_TaxID=4924;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 58785 / CBS 6054;
 RX MEDLINE=96090133; PubMed=7483848;
 RA Hallborn J., Walfriedson M., Penttilae M., Keraenen S.,
 RA Hahn-Hagerd L.B.;
 RT "A short-chain dehydrogenase gene from Pichia stipitis having D-
 arabinitol dehydrogenase activity.";
 RL Yeast 11:839-847(1995).
 CC -1- CATALYTIC ACTIVITY: D-arabinitol + NAD(+) = D-ribulose + NADH.
 CC -1- PATHWAY: D-arabinitol biosynthesis; last step.
 CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
 (SDR) family.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; Z46866; CAA86939.1; -.
 DR PIR; S57351; S57351.
 DR HSSP; Q9ZFY9; 1PK8.
 DR InterPro; IPR002198; ADH_short.
 DR InterPro; IPR002347; Adh_short_C2.
 DR Pfam; PF00106; adh_short_1.
 DR PRINTS; PR00081; GDRDH.
 DR PRINTS; PR00080; SDRPAMILY.
 DR TIGRFAMs; TIGR01830; 3oxo_ACP_reduc; 1.
 DR TIGRFAMs; TIGR01831; fabG_rel; 1.
 DR TIGRFAMs; TIGR01832; kduJ; 1.
 DR PROSITE; PS00061; ADH_SHORT; 1.
 KM NAD; Oxidoreductase.
 FT NP_BIND 22
 FT ACT_SITE 181 181 NAD (By similarity).
 FT BINDING 166 166 Proton acceptor (By similarity).
 FT BINDING 166 166 Substrate (By similarity).
 SQ SEQUENCE 278 AA; 30003 MW; 36869165F23964E6 CRC64;

Query Match 54.3%; Score 798.5; DB 1; Length 278;
 Best Local Similarity 54.3%; Pred. No. 1e-56;
 Matches 151; Conservative 54; Mismatches 64; Indels 9; Gaps 3;

QY 3 DY-----IPTFRPDGHLTVTGACGGLAEALIKGLAYGSDIALIDDOEKTPAKQAQYH 57
 DB 2 DYVYANVVPNFRIDGRLALITIGSGGLAAVSRALAKGADVALLVDMNLEKRTQAARDVL 61
 QY 58 KATEELKKEVPPKGG---STACDISDSDTVHKVPAQVAKDFEKLPLHLVNTAGYGFCEP 114

```

Db      62 GNGEETLKGEHSAICQVANSNCIGDAVDATESSINEHKKIADLLINTAGYCENP 121
Qy      115 CEDYPKAKNAERKVKNNLGSLLVYSGAFAPKLIKEGIGASVVLIGSMGAIYNDPONQV 174
      122 AATYPTANESIMKVGSGSFYVSGFAPLQNNLRG-SITLIGSMGTIYNDPQCM 180
Db      175 YNMSKAGVHLAKTLACCEWAKYNIRNLSNPGYIYPLTKVYINGNEELYNMISGIPQO 234
      181 YNMSKAGVHLVRSILACEWAKYNIRNLTSPGYILPLTRNVI SGHTEKMEKAMESKIPMK 240
Qy      235 RMSEPEYIGAVLYLSESAASYTGASLLVDGGFTSW 272
      241 RMSEPEYIGAVLYLSESAASYTGASLLVDGGFTSW 278
Db

RESULT 4
ASDH_CANTR
ID      ARDH_CANTR      STANDARD;      PRT;      282 AA.
AC      P50166;
DT      01-OCT-1996 (rel. 34, Created)
DT      01-OCT-1996 (rel. 34, Last sequence update)
DT      25-OCT-2004 (rel. 45, Last annotation update)
DE      D-arabinitol 2-dehydrogenase [ribulose forming] (EC 1.1.1.250) (ARDB).
GN      Name:ARD;
OS      Candida tropicalis (Yeast).
OC      Buxaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC      Saccharomycetales; mltosporic Saccharomycetales; Candida.
OX      NCBI_TaxID=54482;
RN      [1]
RP      SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC      STRAIN=ATCC 750;
RX      MEDLINE=95212917; PubMed=7698655; DOI=10.1016/0378-1119(94)00900-D;
RA      Murray J.S., Wong M.L., Miyada C.G., Switchenko A.C., Goodman T.C.,
      Wong B.;
RT      "Isolation, characterization and expression of the gene that encodes
      D-arabinitol dehydrogenase in Candida tropicalis.";
RL      Gene 155:123-128(1995).
RN      [2]
RP      CHARACTERIZATION.
RX      MEDLINE=94071892; PubMed=8250887;
RA      Quong M.W., Miyada C.G., Switchenko A.C., Goodman T.C.;
RT      "Identification, purification, and characterization of a D-arabinitol-
      specific dehydrogenase from Candida tropicalis.";
RL      Biochem. Biophys. Res. Commun. 196:1323-1329(1993).
CC      -1- CATALYTIC ACTIVITY: D-arabinitol + NAD(+) = D-ribulose + NADH.
CC      -1- PATHWAY: D-arabinitol biosynthesis; last step.
CC      -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
      (SDR) family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
      between the Swiss Institute of Bioinformatics and the EMBL outstation -
      the European Bioinformatics Institute. There are no restrictions on its
      use by non-profit institutions as long as its content is in no way
      modified and this statement is not removed. Usage by and for commercial
      entities requires a license agreement (See http://www.isb-sib.ch/announce/
      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; U00675; AAA6355.1; -.
DR      PIR; JC4041; JC4041.
DR      HSSP; Q9ZFY9; 1FK8.
DR      InterPro; IPR002198; ADH_short.
DR      InterPro; IPR002347; Adh_short_C2.
DR      Pfam; PF00106; adh_short_1.
DR      PRINTS; PR00081; GDRDH.
DR      PRINTS; PR00080; SDRFAMILY.
DR      TIGRPFAMs; TIGR01830; 3oxo_ACP_reduc; 1.
DR      TIGRPFAMs; TIGR01831; fadg_tel; 1.
DR      TIGRPFAMs; TIGR01832; kdud; 1.
DR      PROSITE; PS00061; ADH_SHORT_1.
KW      Direct protein sequencing; NAD; Oxidoreductase.
FT      NP_BIND 26 48
FT      ACT_SITE 185 185 NAD (By similarity).
FT      BINDING 170 170 Proton acceptor (By similarity).
FT      BINDING 170 170 Substrate (By similarity).

```

```

SQ      SEQUENCE 282 AA; 30748 MW; A82A3DA4E771EF0D CRC64;
Query Match 56.3%; Score 797; DB 1; Length 282;
Best Local Similarity 55.0%; Pred. No. 1.4e-56;
Matches 149; Conservative 57; Mismatches 61; Indels 4; Gaps 2;
Qy      5 ITPRPDGLTIVTGAACGLAELIKGLIYAGSDIALDIDDKTPAAQAEYHKYATEEL 64
      13 VPSFLDGLVITITGSGGLSAVVRALLAKADIALIDMNERLOQAARVLDQNGEEQM 72
Db      13 KKE---VPMKSYACDISDSDTYHKVPAQVAKDGGKPLHLVNTNAGYENPCEDYPAK 121
      65 KKHESPIGVSAWSCNIGDAEAVELTPVAINEHGVKASVYLINTAGAEENPAEYPAK 132
      73 KKHESPIGVSAWSCNIGDAEAVELTPVAINEHGVKASVYLINTAGAEENPAEYPAK 132
Qy      122 NAEKVKVNLGSLVYSGAFAPKLIKEGIGASVVLIGSMGAIYNDPONQVYNNMSKAG 181
      133 NAEKVKVNLGSLVYSGAFAPKLIKEGIGASVVLIGSMGAIYNDPONQVYNNMSKAG 191
Db      192 VIHLAKTLACCEWAKYNIRNLSNPGYIYPLTKVYINGNEELYNMISGIPQOAKSEPK 241
      192 VIHLAKTLACCEWAKYNIRNLTSPGYILPLTRNVI SGHTEKMEKAMESKIPMKRAEPK 251
Qy      242 YIGAVLYLSESAASYTGASLLVDGGFTSW 272
      252 YIGAVLYLSESAASYTGASLLVDGGFTSW 282
Db

RESULT 5
Q6BY61
ID      Q6BY61      PRELIMINARY;      PRT;      278 AA.
AC      Q6BY61;
DT      25-OCT-2004 (TRENBLrel. 28, Created)
DT      25-OCT-2004 (TRENBLrel. 28, Last sequence update)
DT      25-OCT-2004 (TRENBLrel. 28, Last annotation update)
DE      Similar to CA3288|CAARD8 Candida albicans CAARD8 D-arabinitol
      dehydrogenase.
GN      ORFNames=DEHA0A12595g;
OS      Debaryomyces hansenii CBS767.
OC      Buxaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC      Saccharomycetales; Saccharomycetaceae; Debaryomycetes.
OX      NCBI_TaxID=284592;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=CBS767;
RG      Genolevures;
RA      Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
      Latouche I., de Montigny J., Marck C., Nevegilise C., Talla E.,
      Goffard N., Frangeul L., Aigle M., Anthonard V., Babour A., Barbe V.,
      Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykaeten C.,
      Boissame A., Boyer J., Catolico L., Confanioleri F., de Darvar A.,
      Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
      Hantrege F., Hemequin C., Jauniaux N., Joyet P., Kachouri R.,
      Kerrest A., Kozul R., Lemaire M., Lesur I., Ma L., Muller H.,
      Nicud J.M., Nikolaki M., Oztas S., Ozler-Kalogeropoulos O.,
      Pellens S., Potier S., Richard G.F., Straub M.L., Suleau A.,
      Swennen D., Tekla F., Weselowski-Louvel M., Westhof E., Wirth B.,
      Zenou-Meyer M., Zivanovic I., Bolotin-Pukhara M., Thierry A.,
      Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,
      Wincker P., Souciet J.L.;
      "Genome evolution in yeasts.";
      Nature 430:35-44(2004).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=CBS767;
RA      Genoscope;
RA      Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC      -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
      (SDR) family.
DR      EMBL; CR382133; CAG84833.1; -.
DR      GO; GO:0016491; F:oxidoreductase activity; IEA.
DR      GO; GO:0008152; F:metabolism; IEA.
DR      InterPro; IPR002198; ADH_short.
DR      InterPro; IPR002347; Adh_short_C2.

```


CC (SDR) family.
 DR EMBL; 382132; CNG7697.1; -
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR002347; Adh_short_C2.
 DR InterPro; IPR002347; Adh_short_C2.
 DR InterPro; IPR002347; Adh_short_C2.
 DR Pfam; PF00106; adh_short; 1.
 DR PRINTS; PR00081; GDRHDH.
 DR PRINTS; PR00081; SDRFAMILY.
 DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
 DR Oxidoreductase.
 KW SEQUENCE 343 AA; 37082 MW; 1E593CF047E5585 CRC64;

Query Match 37.3%; Score 528.5; DB 2; Length 343;
 Best Local Similarity 36.9%; Pred. No. 1.3e-34;
 Matches 125; Conservative 55; Mismatches 70; Indels 89; Gaps 10;

QY 16 IVTGACGGLAELIKGLAYGSDIALLD--IDQETAAQ-AEYKATATE----- 62
 DB 12 VITGSGGLGIAAQAQLAAQASVALVDNNLPRIQPAEQLEMYKYNAAHNNVPTPI 71
 QY 63 -----BLKL----- 66
 DB 72 YASPTGKRVSTETESTTGLENSPHDITPDISLSDASDSSQSVAHDAARAHEAG 131
 QY 67 -----KEVP--KMSYACDISD---SPTVAKVPAQVAKDFGKLP-LVNTAGYCENF 113
 DB 132 IPRGKQNPPOORISAMACDVSDVAGVSDTVKAI-----REHNSPDLIDVNCAPFCENM 186
 QY 114 PCEDYPAKAEKMYKYNLIGSLYVSQAAPKPLIKGIGKASVVLIGSMGALVNDPQOV 173
 DB 187 TAFDPNPQVKRLDVLNLMGSSYNFATEVAKSLVLDSPG-SILVASMGSGLVNDPQPT 245
 QY 174 VYNSKAGVYHLAKTLACEMAKYTRVNSINPGYTYGPTKXVINGNEELVNRMTSGIPQ 233
 DB 246 PYNSKAGVYHLAKTLACEMAKYTRVNSINPGYTYGPTKXVINGNEELVNRMTSGIPQ 305
 QY 234 QRMSEPKYIGAVLYLSESASVYTGASLLVDGSGFTSW 272
 DB 306 RMAEPSEFGPIVPMASD-ASSVMTGHDLYDGGYITW 343

RESULT 8

Q6T5L8 PRELIMINARY; PRT; 368 AA.
 AC 06T5L8:
 DT 05-JUL-2004 (TREMBLrel. 27, Created)
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
 DE D-arabinicol dehydrogenase (EC 1.1.1.250).
 GN Name:arbd;
 OS Emericella nidulans (Aspergillus nidulans).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eutrophiales; Trichocomaceae; Emericella.
 OK NCBI_TaxID=162425;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Rodriguez J.M., Ruiz-Sala P., Ugarte M., Penajaya M.A.;
 RL Submitted (OCT-2003) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX PubMed=14612443; DOI=10.1074/jbc.M310055200;
 RA Rodriguez J.M., Ruiz-Sala P., Ugarte M., Penajaya M.A.;
 RT "Pungal Metabolic Model for 3-Methylcrotonyl-CoA Carboxylase
 Deficiency.";
 RL J. Biol. Chem. 279:4578-4587(2004).
 CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
 (SDR) family.
 DR EMBL; AY434691; AAR85467.1; -
 DR EMBL; AY387591; AAR23109.1; -
 DR GO; GO:0047038; F:D-arabinicol 2-dehydrogenase activity; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.

DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR002347; Adh_short_C2.
 DR InterPro; IPR002347; Adh_short_C2.
 DR Pfam; PF00106; adh_short; 1.
 DR PRINTS; PR00081; GDRHDH.
 DR PRINTS; PR00081; SDRFAMILY.
 DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
 DR Oxidoreductase.
 KW SEQUENCE 368 AA; 39686 MW; 67DB6F1CDB44CF62 CRC64;

Query Match 36.1%; Score 510.5; DB 2; Length 368;
 Best Local Similarity 40.3%; Pred. No. 4e-33;
 Matches 106; Conservative 60; Mismatches 94; Indels 3; Gaps 3;

QY 8 FRPDGLTIVGACGGLAELIKGLAYGSDIALLDIDQETAAQAEYKATATELTK 67
 DB 107 FSLERVYLVTLGARGGLVWGQGIYVAGSDLAIDLNHAEBSARKLVQEQENPGLE 166
 QY 68 EYPMKGSYACDISDSPTVAKVPAQVAKDFGKLP-LVNTAGYCENFPCEDYPAKAEKMY 127
 DB 167 QLPNTAHYADVNSPDSVYNSALSEVIAHGKID-HLVTSAGFTENFDAISYPDRMQKLW 225
 QY 128 KYNLLGSLYVSQAAPKPLIKGIGKASVVLIGSMGALVNDPQOVVYNSKAGVYHLAK 187
 DB 226 GNVVDGYTLFATGVAKHMERKAPG-STVMIGSMGALVNDPQAPYNAALAAVRHLAS 284
 QY 188 TLACEMAKYTRVNSINPGYTYGPTKXVINGNEELVNRMTSGIPQRMSEPKYIGAVL 247
 DB 285 SPAYEMAHGIRVNCISGVYLTALTRILDENBELRKWLSLPTGKQVPEDMGAVT 344
 QY 248 YLSESASVYTGASLLVDGSGFT 270
 DB 345 FLTSD-ASKYITGADLRVDGYT 366

RESULT 9

Q7SE47 PRELIMINARY; PRT; 371 AA.
 AC 07SE47:
 DT 01-MAR-2004 (TREMBLrel. 26, Created)
 DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Hypothetical protein.
 GN Name=NCU02128.1;
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
 OK NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OR74A;
 RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
 RA Jaffe D., Fitzhugh W., Ma L.-O., Smirnov S., Purcell S., Reisman B.,
 RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endlitz M.,
 RA Qui D., Iankieff P., Pedersen D., Nelson M., Washburne M.,
 RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zeller A., Schulte U.,
 RA Kothe G.O., Jedd G., Mewes W., Straben C., Marcotte E., Greenberg D.,
 RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnetre S.,
 RA Kanai M., Kamysseis M., Mancini E., Biele C., Rudd S., Fishman D.,
 RA Kryzstofova S., Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S.,
 RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,
 RA Desouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
 RA Varden O., Piamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
 RA Nativg D.O., Alex L.A., Mannhaug G., Ebbole D.J., Freitag M.,
 RA Paulsen I., Sachs M.S., Lander E.S., Nussbaum C., Birren B.;
 RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa."
 CC Nature 0:0-0(2003).
 CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
 (SDR) family.
 DR EMBL; AABX0100038; EAA35050.1; -
 CC -1- CAUTION: The sequence shown here is derived from an
 EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
 preliminary data.

HSSP; P42556, 1P33.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR002198; ADH_short.
 DR InterPro; IPR002347; Adh_short_C2.
 DR Pfam; PF00106; adh_short; 1.
 DR PRINTS; PRO0081; GDRHD.
 DR PRINTS; PRO0080; SDRFAMILY.
 DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
 DR Hypothetical protein; Oxidoreductase.
 KW SEQUENCE 371 AA; 36631 MW; 4866601508P645 CnC64;

Query Match 35.4%; Score 501; DB 2; Length 371;
 Best Local Similarity 40.3%; Pred. No. 2,4e-32;
 Matches 108; Conservative 57; Mismatches 97; Indels 6; Gaps 5;

DB	Accession	Score	Length	Indels	Gaps
QY	5 IPIRFDDHLTVTGACGGLABALIKGLIAYGSDIALIDIDDEKTAQAQAE--YHKYATE	62			
Db	98 LATFSLQKGVGVGTGARGGLGVLMGQGVIVSGSDIAYIDLNKEB-ATROAQNIYETFFKDD	156			
QY	63 ELKLEKVRKMSYADIDSDSDVHVAVPQVAVKDPKELPLHLVNTAGYCENPCEDEYPAKN	122			
Db	157 YPSAKKIKPRVTAHYADVSDPASPVDACIARIYAEHKRID-NLVTSAGFTENFAVYFVDR	215			
QY	123 AEKWKVNLGLSLYVSOAFKPLIKEGIKGASVVLGSMGSAIVNDPQNVVYNNKAGV	182			
Db	216 LRLKMGVAVDGYLFAITAIARLMDRNSG-SLYVIGSMSSGIVAVPQGFYNSKAY	274			
QY	183 IHLAKTLACEMAKNIRVNSLNGYIYGLPLTVNVINGEELYNRMISGIPQQRSEPEKEX	242			
Db	275 RHLASPAVENAKGIRVNCISPGYMLFRLTKKIIDENPDILKRWTSILIPQGMGNPDL	334			
QY	243 IGVAVLYLSESAASYTTGASLLVDGFT	270			
Db	335 MGPVTFLLSD-ASRYVTGADLFVDGXT	361			

RESULT 10
 06C6W8 ID Q6C6W8 PRELIMINARY; PRT; 291 AA.
 AC Q6C6W8;
 DT 25-OCT-2004 (TRENBLREL. 28, Created)
 DT 25-OCT-2004 (TRENBLREL. 28, Last sequence update)
 DT 25-OCT-2004 (TRENBLREL. 28, Last annotation update)
 DE Similar to tci [Q9WVD3 Thermotoga maritima TM0297 Oxidoreductase.
 GN ORFNAME=VALIOE05643g;
 OS Yarrowia lipolytica CLIB99.
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Dipodascaceae; Yarrowia.
 OK NCBI_Taxid=284591;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CLIB99;
 RG Genolayers;
 RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
 RA Lafontaine I., de Montigny J., Marc C., Neveglise C., Talla E.,
 RA Goffard N., Frangoul L., Aigle M., Authourard V., Babour A., Barbe V.,
 RA Barnay S., Blanchard S., Beckerich J.M., Beyne E., Bleykasten C.,
 RA Boisrame A., Boyer J., Cattolico L., Confantieri F., de Daruvar A.,
 RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Gropi A.,
 RA Hantayre F., Hannequin C., Janniaux N., Joyet P., Kachouri R.,
 RA Kerrest A., Kozul R., Lemaire M., Leduc I., Ma L., Muller H.,
 RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
 RA Pellenn S., Potier S., Richard G.F., Straub M.L., Suleau A.,
 RA Swenne D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
 RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
 RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,
 RA Wincker P., Souciet J.L.;
 RT "Genome evolution in yeasts."
 RL Nature 430:35-44 (2004).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CLIB99;

Genoscope;
RA Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
CC (SDR) family.
CC EMBL, CR382131; GAG79175.1; -
DR GO; GO:0016491; F:oxidoreductase activity; ISA.
DR GO; GO:0008152; P:metabolism; ISA.
DR InterPro; IPRO02198; ADH_short.
DR InterPro; IPRO02347; Adh_short_C2.
DR Pfam; PF00106; adh_short; 1.
DR PRINTS; PR00081; GDRDH.
DR PRINTS; PR00080; SDRFAMILY.
DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
KW Oxidoreductase.
SQ SEQUENCE 291 AA; 31002 MW; 9479873014C4101D CRC64;

Query Match 32.9%; Score 465; DB 2; Length 291;
Best Local Similarity 36.9%; Pred. No. 1.5e-29;
Matches 103; Conservative 55; Mismatches 85; Indels 36; Gaps 7;

QY DYITPFPDHLITVTGACCGLAELIKGLAYG-SDIALDLIDOE-----KTA 51
DB 36 NPLSKFRIDGVALVITGANGLGFSMAEGSCVSGKIALTLVDQDGLDAIEKIKAYG 95
QY 52 KQAEYHKYATEELTKAEVPMKGSYACDISDPTVHKVFAVADFGLPLHLVNTGYCE 111
DB 96 VQQAIFYK-----ADVDEESVNEIIDRVHDLGSVDV-VNNSAGVAD 136
QY 112 NFPCEPDYPAKNAEMKVKVNLGLSYQAAPKPLIEGIGASVYLIGMSGAVLNDPON 172
DB 137 LVHAAYEPADFRFVRIIDININGSFLVQAARHNTKQG-TGGTVVFLASVSGSIVNWP 195
QY 172 QVVNMSGAGYTHLAKTLACGMAKXNRVNSLNPNGYIGPLTKVINGNEELYNRWISGI 231
DB 196 QSAVNASGAAYHLSKSLSLAEMAVHNRKMSISISGYDPTLNR-----AVNTLEEMWDR 251
QY 232 PQORMSEPKYIGAVLYLLSESASVTTGASLVDGFT 270
DB 252 PLGRUGDPDELTGACITLADP-ASVYTTGSDIITDGYT 289

RESULT 11
Q6SSPI
ID Q6SSPI PRELIMINARY; PRT, 259 AA.
DT 25-OCT-2004 (TREMBlrel. 28. Created)
DT 25-OCT-2004 (TREMBlrel. 28. Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28. Last annotation update)
DE Fabg protein.
GN Name=fabg; ORFNames=MS1412;
OS Mannheimia succiniciproducens MBEL55R.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Mannheimia.
OX NCBI_TaxID=221988;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=MBEL55F.
RA Hong S.H., Kim J.S., Lee S.Y., In Y.H., Choi S.S., Rih J.-K.,
RA Kim C.H., Jeong H., Hur C.-G., Kim U.G.,
RA "The genome sequence of the capnophilic rumen bacterium Mannheimia
RT succiniciproducens".
RL Nat. Biotechnol. 0:0-0(2004).
DR EMBL, AB016827; AAU38019.1; -
SQ SEQUENCE 259 AA; 27552 MW; B46EFA78B3F5923C CRC64;

Query Match 28.8%; Score 407; DB 2; Length 259;
Best Local Similarity 37.4%; Pred. No. 6.7e-25;
Matches 99; Conservative 44; Mismatches 104; Indels 18; Gaps 6;

QY 9 RPDGHLITVTGACCGLAELIKGLAYGSDIALDLIDP---QEKPAKQAEYHKYATEELK 65
DB 8 KLTGKTAPVTGARGIGKSVLAIFAQAGANVTIADFIAAEAKTAA-----ETIA 56

QY 66 LKEVPKMGSYACDISDSDTVHKVFAQVAKDPEGLPLHLVNTAGCENPCEDYPAKNAEK 125
 DB 57 KEEGVKASJAVOTVDVDOASVNHLMVDVTKQFGLKDIACFN-AGICTNPAEEMSYQWMLK 115
 QY 126 MKVKNVLSGLSYVQAFAKPLIKEGIKGASVVLIGSMGAIIVNDPONOVYNNKSKAGVHL 185
 DB 116 VINVNLGVPLFAQAAGKLMIEGQ-TGGSIIINTASNAHVVPQPCANVASKAGVLOL 174
 QY 186 AKTLACBAKXNIRVNSLNPGLTYPLTKVNVINGNEELYNRMISGIPQORMSEPKRYTGA 245
 DB 175 TKSLAIEMAKNIRVNSLSPGYIGTELTLN-SKDQPLIKEMVNAAPLHLRQCPRELOSI 233
 QY 246 VLYLSESASAYTGSASLLVDCGFT 270
 DB 234 CVYLAGDT-SSPTTGADPTVDGAFT 257

RESULT 12

Q9WYD3 PRELIMINARY; PRT; 257 AA.
 ID Q9WYD3
 AC Q9WYD3; 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 12, Last annotation update)
 DE Oxidoreductase, short chain dehydrogenase/reductase family.
 GN OrderedLocustNames=TM0297;
 OS Thermotoga maritima.
 OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
 OX NCBI_TaxID=2336;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSB8 / DSM 3109 / ATCC 43589;
 RX MEDLINE=99287316; PubMed=10360571; DOI=10.1038/20601;
 RA Neilson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,
 RA Hat D.H., Hickey B.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
 RA McDonald L.A., Utterback T.R., Malek J.A., Liner K.D., Garrett M.M.,
 RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.L.,
 RA Heidelberg J.F., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
 RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from
 RT genome sequence of Thermotoga maritima";
 RL Nature 399:323-329(1999).
 CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
 CC (SDR) family.
 CC EMBL: AB001711; AAD35385.1; -
 DR PIR: A72395; A72395.
 DR HSSP: Q9ZEY9; 1FTH.
 DR TIGR: TM0297;
 DR GO: GO:0016491; P:oxidoreductase activity; IEA.
 DR GO: GO:0008152; P:metabolism; IEA.
 DR InterPro: IPR002198; ADH_short.
 DR InterPro: IPR002347; Adh_short_C2.
 DR Pfam: PF00106; adh_short; 1.
 DR PRINTS: PRO0081; GDRDH.
 DR PROSITE: PS00080; SDRFAMILY.
 DR PROSITE: PS00061; ADH_SHORT; UNKNOWN_1.
 DR Complete proteome; Oxidoreductase.
 KM SEQUENCE 257 AA; 27886 MW; 724484F3EDF5EAC CRC64;

Query Match 28.4%; Score 402.5; DB 2; Length 257;
 Best Local Similarity 36.9%; Pred. No. 1.5e-24;
 Matches 101; Conservative 50; Mismatches 90; Indels 33; Gaps 11;

QY 8 FRPFGHLLTVYAGCGLAELIKGLAYGSDIALDDIOEKTAQAEYHKAATBELK 67
 DB 8 FSLKRAVAVTGGGCGIGKALAAAGAVLIMDNEB-TARKTYE----- 54
 QY 68 EVPRKGS---YACDISDSDTVHKVFAQVAKDPEGLPLHLVNTAG---YCNFPCEDYPA 120
 DB 55 EIKKSGEADRYVDVTKQFGLKDIACFN-AGICTNPAEEMSYQWMLK 110
 QY 121 KNAEKMKVKNVLSGLSYVQAFAKPLIKEGIKGASVVLIGSMGAIIVNDPONOVYNNKSKA 180

DB 111 EKKKKVIVDNLVG-VFLSAKAEFLAMKER-KYCKTINIANSNGHIVNKPQKOTAVNASKA 168
 QY 181 GVHILAKTACEMAKYINRVNSLNPGLTYPLTKVNVINGNEELYNRMISGIPQORMSE 238
 DB 169 GVHILTRSLAEMAPYGRVNSISRGYIRTLISPNV-----KDLVPLMLMIPILGRIGE 224
 QY 239 PREYIGAVVLYLSESASAYTGSASLLVDCGFTSW 272
 DB 225 VDDLGAALIFLAS-PASDYMTHGHDVLDGYTVM 257

RESULT 13

Q6BN16 PRELIMINARY; PRT; 291 AA.
 ID Q6BN16
 AC Q6BN16; 25-OCT-2004 (TREMBlrel. 28, Created)
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
 DE Similar to YAL10B16192g Yarrowia lipolytica.
 DE ORFNames=DEHA0F01177g;
 OS Debaryomyces hanseni CBS767.
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetaceae; Saccharomycetaceae; Debaryomycetes.
 OX NCBI_TaxID=284592;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CBS767;
 RG GenomLevures;
 RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S., Talla E.,
 RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,
 RA Gottard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
 RA Barney S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
 RA Bolarine A., Boyer J., Cattolico L., Confiantieri F., de Daruvar A.,
 RA Despoix L., Fabre B., Fairhead C., Ferry-Dumazet H., Groppi A.,
 RA Hantrepe F., Hemequin C., Janniaux N., Joyet P., Kachouri R.,
 RA Kerrest A., Kozul R., Lemaire M., Lesur I., Ma L., Muller H.,
 RA Nicaud J.M., Nikolaki M., Oztas S., Ozler-Kalogeropoulos O.,
 RA Pellenn S., Potier S., Richard G.F., Straub M.L., Suleau A.,
 RA Swenne D., Tekala F., Mesolowski-Louvel M., Westhof B., Wirth B.,
 RA Zenon-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
 RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weisenbach J.,
 RA Wincker P., Souciet J.L.;
 RT "Genome evolution in yeasts";
 RL Nature 430:35-44(2004).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CBS767;
 RA Genoscope;
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
 CC (SDR) family.
 CC EMBL: CR382138; CAG88708.1; -
 DR GO: GO:0016491; P:oxidoreductase activity; IEA.
 DR GO: GO:0008152; P:metabolism; IEA.
 DR InterPro: IPR002198; ADH_short.
 DR InterPro: IPR002347; Adh_short_C2.
 DR Pfam: PF00106; adh_short; 1.
 DR PRINTS: PRO0081; GDRDH.
 DR PROSITE: PS00080; SDRFAMILY.
 DR PROSITE: PS00061; ADH_SHORT; UNKNOWN_1.
 DR Oxidoreductase.
 KM SEQUENCE 291 AA; 31685 MW; 7386F67E7DDBE9BD CRC64;

Query Match 28.2%; Score 399.5; DB 2; Length 291;
 Best Local Similarity 36.8%; Pred. No. 3.2e-24;
 Matches 100; Conservative 50; Mismatches 96; Indels 27; Gaps 10;

QY 5 IPRFPGHLLTVYAGCGLAELIKGLAYGSDIALDDIOEKTAQAEYHKAATBELK 64
 DB 35 IDLFNLKRAVAVTGGGCGIGKALAAAGAVLIMDNEB-TARKTYE----- 85
 QY 65 KNAEKMKVKNVLSGLSYVQAFAKPLIKEGIKGASVVLIGSMGAIIVNDPONOVYNNKSKA 180

Db 86 KLVGV-KLRSYCVDTDSKEVQTVNNIVEDFDIDIFVAN-AGI--NIPVGIINEENS 141
QY 119 PAKNAEKMYKVNILGSLYVSOAFKPLIK--EGIKASVVLIGMSGALVNDPQNVVN 176
Db 142 NDKRMVLVMDINILNGVYCAKNGKVSRRKGRG-SMITGMSGHIIINTPVHQAAYN 200
QY 177 MSKAGVHLKLACCEMAKNIRVNSLNPYIYGPLTKVINGNEELYNRMISGIPQORM 236
Db 201 ASKASVHFPAKSLIEVDVF-ARVNSISPGYVNSGINDHLPSTNRK---RWSSTIPMGRE 256
QY 237 SEPKYIGAVLYLSESASVYTTGASLLVDGFF 269
Db 257 GLPELVGAVLYLASD-ASTYTTADIIIDGCG 288
RESULT 14
DCXR_TRIPE STANDARD; PRT; 266 AA.
ID DCXR_TRIPE STANDARD; PRT; 266 AA.
AC O8WK50;
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DE 25-OCT-2004 (Rel. 45, Last annotation update)
DE L-xylulose reductase (EC 1.1.1.10) (XR).
GN Name=Lxrl;
OS Trichoderma reesei (Hypocrea jecorina).
OC Burkholderia; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.
OX NCBI_TaxID=51453;
RN [1]
RP SEQUENCE FROM N.A., AND ENZYME ACTIVITY.
RX MEDLINE=22005794; PubMed=12009906; DOI=10.1021/bi025529i;
RA Richard P., Putkonen M., Vasaenen R., Londestorouh J.,
RA Penttilae M.,
RT "The missing link in the fungal L-arabinose catabolic pathway,
RT identification of the L-xylulose reductase gene."
RL Biochemistry 41:6432-6437(2002).
CC -I- FUNCTION: Catalyzes the NADPH-dependent reduction of L-xylulose,
CC D-xylulose, D-fructose, and L-sorbose, with the highest affinity
CC for L-xylulose.
CC -I- CATALYTIC ACTIVITY: xylitol + NADP(+) = L-xylulose + NADPH.
CC -I- PATHWAY: L-arabinose pathway; third step.
CC -I- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
CC (SDR) family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AF375616; AAM20896.1; -.
DR HSSP; Q9ZFY9; 1FK8.
DR InterPro; IPR002198; ADH_short.
DR InterPro; IPR002347; Adh_short_C2.
DR Pfam; PF00106; adh_short; 1.
DR PRINTS; PR00081; GDRHDH.
DR PRINTS; PR00080; SDRFAMILY.
DR TIGRFAMS; TIGR01830; 3oxo_ACP_reduc; 1.
DR TIGRFAMS; TIGR01831; fabG_reli; 1.
DR TIGRFAMS; TIGR01832; kdud_1.
DR PROSITE; PS00061; ADH_SHORT; PALSE_NEG.
KW Carbohydrate metabolism; NADP; Oxidoreductase; Xylulose metabolism.
KW NP BIND 23
FT ACT SITE 174 174 NADP (By similarity).
FT ACT SITE 178 178 Proton acceptor (By similarity).
FT BINDING 159 159 By similarity.
FT BINDING 159 159 Substrate (By similarity).
SQ SEQUENCE 266 AA; 28478 MW; 1CF56334DA86F109 CRC64;
Query Match 27.8%; Score 393; DB 1; Length 266;
Best Local Similarity 36.4%; Pred. No. 9, 6e-24;
Matches 102; Conservative 51; Mismatches 103; Indels 24; Gaps 8;

QY 1 MTDYIPT-----FRPDGHLTVTGACG--GLAELIKGLAYSGDIALLDIDQEKTA 51
Db 1 MPQPVPTARLNLDFSLKGVVVVVTGASGPRGKIGIBAAGCAMEGADLAITYSSRKEGAE 60
QY 52 KQAEHYKATVEELKKEVPMGMSYACDISPDVTHKVFPAQVADPFCKLPLHLVNTGYCE 111
Db 61 KNA-----EELTKREVGVVKKVYKQSDYNDVERFVQVQVSDFFKIDAFIYN-AGATA 112
QY 112 NFPCEYPAKNAEKMYKVNILGSLYVSOAFKPLIKEGIKASVVLIGMSGALVNDPQN 171
Db 113 NSGVDSASDMDVHVQVDSLGTAICAKVGAHFKKQ--HSLVLTATASMGVANYPOE 170
QY 172 QVYNNMSKAGVHLAKTACCEMAKNIRVNSLNPYIYGPLTKVINGNEELYNRMISG 231
Db 171 QTSYVNAKAGCCHLASLANEMRDF-ARVNSISPGYIDTGLSDFIDEKTOEL---WRSMI 226
QY 232 PQORMSEPKYIGAVLYLSESASVYTTGASLLVDGFFS 271
Db 227 PMGRNGDAKELKGAAYVLVSD-ASSYTTGADIVIDGGYTT 265
RESULT 15
Q89QA5 PRELIMINARY; PRT; 257 AA.
ID Q89QA5 PRELIMINARY; PRT; 257 AA.
AC Q89QA5;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Oxidoreductase.
GN OrderedLocustNames=d1j3225;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=USD110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneo T., Nakamura Y., Sato S., Minamisawa K., Uchiyumi T.,
RA Saramoto S., Watanabe A., Idegawa K., Iriuchi M., Kawashima K.,
RA Kohara W., Matsuno M., Shimpō S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USD110."
RL DNA Res. 9:189-197(2002).
CC -I- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
CC (SDR) family.
CC EMBL; AP005946; BAC48490.1; -.
DR HSSP; AP005946; F:oxidoreductase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002198; ADH_short.
DR InterPro; IPR002347; Adh_short_C2.
DR Pfam; PF00106; adh_short; 1.
DR PRINTS; PR00081; GDRHDH.
DR PRINTS; PR00080; SDRFAMILY.
DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
KW Complete proteome; Oxidoreductase.
SQ SEQUENCE 257 AA; 27403 MW; A49BE3BF7F21363D CRC64;

Query Match 27.7%; Score 392.5; DB 2; Length 257;
Best Local Similarity 36.1%; Pred. No. 1e-23;
Matches 99; Conservative 45; Mismatches 107; Indels 23; Gaps 8;
QY 4 YIFRFPDGHLLTVTGACGGLAELIKGLAYSGDIALLDIDQEKTAQKQAEHYKATVEE 63
Db 2 YLEKPKLDKQKAFITGGQGIGLACAEALABAGRVVIGDSDKVAQDAKA-----S 53
QY 64 LKLLK-----EVPKMGSYACDISPDVTHKVFPAQVADPFCKLPLHLVNTGYCEN--FPCEY 118
Db 54 LKAGKFDALYIM-----DVTITRRAVAVANDVARRKGVDI-LVNNAGIGARSETPATV 107

Qy 119 PAKNAERKVKVNLGSLVYSOAFAPKLIKEGIGKASVYLIGSMGAIYNDPONOVYNNMS 178
 Db 108 TDEHMLNVIIDVNLNGTFWCCREFGKMLK-AQSGAIVVGSMSGFIVNKPQEQCFYNAS 165
 Qy 179 KAGVHLAKTLACENAKYINIRVNSINPGYIYGPITKENVINGNEELYNRWISGIPOQRMSB 238
 Db 166 KAGVHLTKSLAERKARGIRVNAVPTIETPLNA-FVKSINPKYIDAWIGTPMARMQ 224
 Qy 239 PREYIGAVLYLISESASVYTTGASLLVDGFTSW 272
 Db 225 VEERIASVVLPLSSB-AASIMTGSIVLVDGTYTCW 257

Search completed: July 9, 2005, 10:30:58
 Job time : 181 secs

This Page Blank (uspio)